

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2005, 09:06:37 ; Search time 44 Seconds
(without alignments)
1812.811 Million cell updates/sec

Title: US-10-696-639-39
Perfect score: 4369
Sequence: 1 MGLPRGPLASLLQLQVCWLQ.....NEWGSRFKKLADMYGGEDD 829
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4369	100.0	829	1 IJHUCP	cadherin 3 precurs
2	3555.5	81.4	822	1 IJMSCP	P-cadherin precurs
3	2573.5	58.9	732	1 IJHUCB	B-cadherin precurs
4	2420	55.4	882	1 IJHUCE	cadherin 1 precurs
5	2388	54.7	884	1 IJMSCE	E-cadherin precurs
6	2379	54.5	884	2 S34438	uvomorulin - mouse
7	2375	54.4	895	1 IJXLCP	EP-cadherin precurs
8	2371	54.3	905	2 S43064	cadherin - African
9	2323	53.2	491	1 IJBOCP	P-cadherin - bovin
10	2274	52.0	887	1 IJHCLL	E-cadherin precurs
11	2192.5	50.2	871	2 S47518	cadherin - African
12	1707	39.1	906	1 IJHUCN	cadherin 2 precurs
13	1704.5	39.0	912	1 IJHCHN	N-cadherin precurs
14	1697.5	38.9	906	1 IJMSCN	N-cadherin precurs
15	1694	38.8	913	1 IJHCKR	R-cadherin precurs
16	1691	38.7	913	1 IJBOCN	N-cadherin precurs
17	1655	37.9	913	1 A47543	R-cadherin precurs
18	1621.5	37.1	916	2 C38992	N-cadherin 4 precurs
19	1621	37.1	783	2 I50116	N-cadherin 2 precu
20	1616.5	37.0	906	1 IJXLC2	N-cadherin 1 precu
21	1609	36.8	905	1 IJXLC1	cadherin-15 precu
22	1321.5	30.2	814	2 G02878	M-cadherin - mouse
23	1302.5	29.8	730	1 IJMSCM	cadherin 13 precu
24	1021.5	23.4	713	2 B38992	cadherin-7 - chick
25	1007.5	23.1	785	2 I50180	cadherin-6B - chic
26	990	22.7	790	2 I50178	cadherin-6 - human
27	982	22.5	790	2 I37016	cadherin-14 - huma
28	976.5	22.4	790	2 G02678	T-cadherin precurs
29	976	22.3	712	1 IJMSCT	

RESULT 1

IJHUCP

cadherin 3 precursor - human
N;Alternate names: P-cadherin; Placental cadherin

C;Species: Homo sapiens (man)
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004

C;Accession: A33659

R;Shimoyama, Y.; Yoshida, T.; Terada, M.; Shimosato, Y.; Abe, O.; Hirohashi, S.

J. Cell Biol. 109, 1787-1794, 1999

A;Title: Molecular cloning of a human Ca(2+)-dependent cell-cell adhesion molecule homol

A;Reference number: A33659; MUID:90009051; PMID:2793940

A;Accession: A33659

A;Molecule type: mRNA

A;Residues: 1-829 <SH1>

A;Cross-references: UNIPROT:P22223; UNIPARC:UPI000004C15E; GB:X63629; NID:G35322; PIDN:C

C;Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought t

C;Genetics:

A;Gene: GDB:CDH3

A;Cross-references: GDB:132860; OMIM:114021

A;Map position: 16q24.1-16qter

C;Superfamily: cadherin; cadherin repeat homology

F;1-26/Domain: signal sequence #status predicted <SIG>

F;27-107/Domain: propeptide #status predicted <PRO>

F;108-829/Product: P-cadherin #status predicted <MAT>

F;108-650/Domain: extracellular #status predicted <EXT>

F;110-215/Domain: cadherin repeat homology <CR1>

F;185-190/Region: cadherin binding #status predicted

F;218-328/Domain: cadherin repeat homology <CR2>

F;331-440/Domain: cadherin repeat homology <CR3>

F;441-548/Domain: cadherin repeat homology <CR4>

F;549-652/Domain: cadherin repeat homology <CR5>

F;653-677/Domain: transmembrane #status predicted <TM>

F;678-829/Domain: intracellular #status predicted <INT>

F;785-800/Region: serine-rich

F;200,566/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 4369; DB 1; Length 829;

Best Local Similarity 100.0%; Pred. No. 1.6e-270; Indels 0; Gaps 0;

Matches 829; Conservative 0; Mismatches 0;

QY 1 MGLPRGPLASLLQLQVCWLQCAASEPCRAVFREAEVTLAAGAEQEGQALGVFMGCPG 60

DB 1 MGLPRGPLASLLQLQVCWLQCAASEPCRAVFREAEVTLAAGAEQEGQALGVFMGCPG 60

QY 61 QEPALFSTNDNDFTVRNGETVQERRSLKERNPLKIPFSGKILRRHKRDWVVAISVPENG 120

DB 61 QEPALFSTNDNDFTVRNGETVQERRSLKERNPLKIPFSGKILRRHKRDWVVAISVPENG 120

QY 121 KGPPORLNQKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDREBIAK 180

DB 121 KGPPORLNQKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDREBIAK 180

```
181 YELFGHAVSENGASVEDPMNISIIIVTDNDHKPKFTODTFRGSLVGLVPGTSMQVMTAT 240
181 YELFGHAVSENGASVEDPMNISIIIVTDNDHKPKFTODTFRGSLVGLVPGTSMQVMTAT 240
241 DEDDAIYTYNGVAVSIHSQEPKDPHDLMTIHRSTGTISVSSGLDREKVPYTLTIQA 300
241 DEDDAIYTYNGVAVSIHSQEPKDPHDLMTIHRSTGTISVSSGLDREKVPYTLTIQA 300
301 TMDGSGSTTTAVAVVEILDANDNAPMFDPKYEAHVPENAVGEVORLVTDLDPNSP 360
301 TMDGSGSTTTAVAVVEILDANDNAPMFDPKYEAHVPENAVGEVORLVTDLDPNSP 360
361 AWRATYLMGGDDGDHFTIITHPESNOGILTTKGLDPEAKNOHTLVVETNAPFVLKL 420
361 AWRATYLMGGDDGDHFTIITHPESNOGILTTKGLDPEAKNOHTLVVETNAPFVLKL 420
421 PTSTATIVVHVEDVNEAPVFPVPSKVVEVOEGIPTEGPEVCVYTAEDPDKENQISYRILR 480
421 PTSTATIVVHVEDVNEAPVFPVPSKVVEVOEGIPTEGPEVCVYTAEDPDKENQISYRILR 480
481 DPAGWLAMDPSGQVTAAGTLDREDEQFVRNNIYEVNVLAMDNGSPPTTGTGLTLTLID 540
481 DPAGWLAMDPSGQVTAAGTLDREDEQFVRNNIYEVNVLAMDNGSPPTTGTGLTLTLID 540
541 VNDHGPVPEPQITICNOSPVRHVNLITDKDLSPTSPPFOAQLTDDSDIYKTAEVNEEGD 600
541 VNDHGPVPEPQITICNOSPVRHVNLITDKDLSPTSPPFOAQLTDDSDIYKTAEVNEEGD 600
601 TVVLSLKKFLKQDYDVHLSLSHGNKEQLTVIRATVCDCHGVETCPGPKWGGFLLPVL 660
601 TVVLSLKKFLKQDYDVHLSLSHGNKEQLTVIRATVCDCHGVETCPGPKWGGFLLPVL 660
661 GAVLALLFLLVLLVLLVLRKKRKIKEPILLPDDTRDNVYFYGGEGGEDDQDYDITOLHR 720
661 GAVLALLFLLVLLVLLVLRKKRKIKEPILLPDDTRDNVYFYGGEGGEDDQDYDITOLHR 720
721 GLEARPEVLNRNDVAPTIIPPMYRPRPANDPEIGNFIENLKAANTDPTAPPYDVLV 780
721 GLEARPEVLNRNDVAPTIIPPMYRPRPANDPEIGNFIENLKAANTDPTAPPYDVLV 780
781 DYEGSGSDAASLSLTSASDQDQDYDLYNEWGSRRFKKLADMYGGGEDD 829
781 DYEGSGSDAASLSLTSASDQDQDYDLYNEWGSRRFKKLADMYGGGEDD 829

RESULT 2
IJMSCP
P-cadherin precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C;Accession: S03163; S34458
R;Nose, A.; Nagafuchi, A.; Takeichi, M.
EMBO J. 6, 3655-3661, 1987
A;Title: Isolation of placental cadherin cDNA: identification of a novel gene family of
A;Reference number: S03163; MUID:88111554; PMID:3428270
A;Accession: S03163
A;Molecule type: mRNA
A;Residues: 1-822 <NOS>
A;Cross-references: UNIPROT:P10287; UNIPARC:UPI000017435B; EMBL:X06340
R;Faraldo, M.L.M.; Cano, A.
J. Mol. Biol. 231, 935-941, 1993
A;Title: The 5' flanking sequences of the mouse P-cadherin gene. Homologies to 5' sequen
A;Reference number: S34458; MUID:93294853; PMID:8515462
A;Accession: S34458
A;Molecule type: DNA
A;Residues: 1-55 <PAR>
A;Cross-references: UNIPARC:UPI000016CCF0; EMBL:X68057
C;Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought t
C;Genetics:
A;Intons: 16/3
C;Superfamily: cadherin; cadherin repeat homology
C;Keywords: calcium binding; cell adhesion; duplication; glycoprotein; placenta; transme
```

```
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-99/Domain: propeptide #status predicted <PRO>
F;100-822/Product: P-cadherin #status predicted <MAT>
F;100-645/Domain: extracellular #status predicted <EXT>
F;102-207/Domain: cadherin repeat homology <CR1>
F;210-320/Domain: cadherin repeat homology <CR2>
F;323-432/Domain: cadherin repeat homology <CR3>
F;433-540/Domain: cadherin repeat homology <CR4>
F;541-645/Domain: cadherin repeat homology <CR5>
F;646-670/Domain: transmembrane #status predicted <TM>
F;671-822/Domain: intracellular #status predicted <INT>
F;778-793/Region: serine-rich
F;192-558/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 81.4%; Score 3555.5; DB 1; Length 822;
Best Local Similarity 81.6%; Pred. No. 1.2e-218; Indels 13; Gaps 5;
Matches 679; Conservative 57; Mismatches 83;

QY 1 MGLPRGPLA-SLLLLQVLCWLQCAASEPCRAVF-REAEVTLLEAGBAEQEPQALGKVMGC 58
DB 1 MELLSGPHAFLLLLQVLCWLSVSPYRAGFIGEAGVTLEVEGTDLEPSQVLGKVALAG 60
QY 59 PQQEPALFSTNDNDFVRNGETVQERRSLKERNPLKIFPSKRLRRHKRWVAVPSVPE 118
DB 61 QGMHHA---DNGDIIIMLTRGTQVGKDAHMS-----PPTRILRRRKREWMVPPIFVPE 110
QY 119 NGKGPORLNOLKSNKRDTKIFYSTIGPGADSPPEGVFAVEKETGWLNLNKLPLDREI 178
DB 111 NGKGPORLNOLKSNKRDGKIFYSTIGPGADSPPEGVFTIEKESGWLNLNKLPLDREI 170
QY 179 AKYELFGHAVSENGASVEDPMNISIIIVTDNDHKPKFTODTFRGSLVGLVPGTSMQV 238
DB 171 VKYELFGHAVSENGASVEDPMNISIIIVTDNDNKPFTODTFRGSLVGLVPGTSMQV 230
QY 239 ATDEDDAIYTYNGVAVSIHSQEPKDPHDLMTIHRSTGTISVSSGLDREKVPYTLTI 298
DB 231 ATDEDDAVNTYNGVAVSIHSQEPKDPHDLMTIHRSTGTISVSSGLDREKVPYTLTI 290
QY 299 QATMDGSGSTTTAVAVVEILDANDNAPMFDPKYEAHVPENAVGEVORLVTDLDPN 358
DB 291 QATMDGSGSTTTAVAVVEILDANDNAPMFDPKYEAHVPENAVGEVORLVTDLDPN 350
QY 359 SPAWRATYLMGGDDGDHFTIITHPESNOGILTTKGLDPEAKNOHTLVVETNAPFVL 418
DB 351 WPAWRATYLMGGDDGDHFTIITHPETNQGVLLTTKGLDPEAQDQHTLVVETNAPFV 410
QY 419 KLPTSTATIVVHVEDVNEAPVFPVPSKVVEVOEGIPTEGPEVCVYTAEDPDKENQISYRI 478
DB 411 KLPTATATVYVHVVDVNEAPVFPVPSKVVEVOEGISIGELVCITYTAQDPDKEDOKISYTI 470
QY 479 LRDPAGWLAMDPSGQVTAAGTLDREDEQFVRNNIYEVNVLAMDNGSPPTTGTGLTLTL 538
DB 471 SRDPANWLAVDPSGQITAAAGILDREDEQFVRNNIYEVNVLAMDNGSPPTTGTGLTLTL 530
QY 539 IDVNDHGPVPEPQITICNOSPVRHVNLITDKDLSPTSPPFOAQLTDDSDIYKTAEVNE 598
DB 531 TDINDHGPVPEPQITICNOSPVPQVNLITDKDLSPNSSPFOAQLTHDSDIYMAVSEK 590
QY 599 GDTVVLSSKFLKQDYDVHLSLSHGNKEQLTVIRATVCDCHGV-ETCPGPKWGGFLL 657
DB 591 GDTVALSLKKFLKQDYDVHLSLSHGNREQLTMIRATVCDCHGVFNDPFRPKWGGFLL 650
QY 658 PVLGAVLALLFLLVLLVLLVLRKKRKIKEPILLPDDTRDNVYFYGGEGGEDDQDYDITQ 717
DB 651 PILGAVLALLFLLVLLVLLVLRKKRKIKEPILLPDDTRDNVYFYGGEGGEDDQDYDITQ 710
QY 718 LHRGLEARPEVLNRNDVAPTIIPPMYRPRPANDPEIGNFIENLKAANTDPTAPPYDVL 777
DB 711 LHRGLEARPEVLNRNDVAPTIIPPMYRPRPANDPEIGNFIENLKAANTDPTAPPYDVL 770
QY 778 LVFDYEGSGSDAASLSLTSASDQDQDYDLYNEWGSRRFKKLADMYGGGEDD 829
DB 771 MVFDYEGSGSDAASLSLTSASDQDQDYDLYNEWGSRRFKKLADMYGGGEDD 822
```

RESULT 3

IJCCHB

B-cadherin precursor - chicken (fragment)
N/Alternate names: K-CAM protein
C/Species: Gallus gallus (chicken)
C/Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 09-Jul-2004
C/Accession: A41634; A38715; S16160
R/Sorkin, B.C.; Gallin, W.J.; Edelman, G.M.; Cunningham, B.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 11545-11549, 1991
A/Title: Genes for two calcium-dependent cell adhesion molecules have similar structures
A/Reference number: A41634; MUID:92107987; PMID:1763068
A/Accession: A41634
A/Molecule type: DNA
A/Residues: 1-732 <SOR>
A/Cross-references: UNIPROT:P33145; UNIPARC:UPI0000136DBF; GB:M81894; NID:g212226; PIDN:
R/Napolitano, E.W.; Venstrom, K.; Wheeler, E.F.; Reichardt, L.F.
J. Cell Biol. 113, 893-905, 1991
A/Title: Molecular cloning and characterization of B-cadherin, a novel chick cadherin.
A/Reference number: A38715; MUID:91225083; PMID:2026653
A/Accession: A38715
A/Molecule type: mRNA
A/Residues: 7-413, 'V', 415-732 <NAP>
A/Cross-references: UNIPARC:UPI00001712B4; GB:X58518; NID:G63113; PIDN:CAA41408.1; PID:9
C/Cross-references: UNIPARC:UPI00001712B4; GB:X58518; NID:G63113; PIDN:CAA41408.1; PID:9
C/Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought b
C/Genetics:
A/Gene: K-CAM
A/Introns: 29/3; 81/3; 130/1; 188/3; 231/3; 293/3; 375/2; 423/1; 498/1; 571/1; 614/3; 66
C/Superfamily: cadherin; cadherin repeat homology
C/Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro
F/1-6/Domain: propeptide (fragment) #status predicted <PRO>
F/6-554/Domain: extracellular #status predicted <EXT>
F/7-732/Product: B-cadherin #status predicted <MAP>
F/9-114/Domain: cadherin repeat homology <CR1>
F/84-89/Region: cadherin binding #status predicted
F/117-227/Domain: cadherin repeat homology <CR2>
F/230-339/Domain: cadherin repeat homology <CR3>
F/340-447/Domain: cadherin repeat homology <CR4>
F/448-552/Domain: cadherin repeat homology <CR5>
F/555-580/Domain: transmembrane #status predicted <TM>
F/581-732/Domain: intracellular #status predicted <INT>
F/689-702/Region: serine-rich
F/137,410/Binding site: carboxydrate (Asn) (covalent) #status predicted

Query Match 58.9%; Score 2573.5; DB 1; Length 732;
Best Local Similarity 66.2%; Pred. No. 4e-156;
Matches 485; Conservative 103; Mismatches 138; Indels 7; Gaps 5;

QY 102 LRRHKRDWVAPISVPENGKGPFPQRLNQLKSKNRDRTKIFYISITGPGADSPPEGVFAVE 161
DB 1 LRQKRDWVPPKVPENEGPFPKNLVQIKSNRDREAKIFYISITGGADAPPEGIFTIE 60
QY 162 KETGWLKLPDRRIAKYELFGHVSNGASVEDPMNISIIVTDNDHKPKPTQDTR 221
DB 61 KETGWNKVTQPLDREHINKYLYSHAVSNGKPEPMEIIVTVDNDNKPQFTQEVER 120
QY 222 GSVLEGLPCTSNVMTATDEDAIYTYNGVAVSIHSGPKDPHDLMTIHRSTCTISV 281
DB 121 GSVPEGALPGTSVMRVNATDADDVETNGVAVSIHSGPEPHPMFTVNRATGLSV 180
QY 282 ISSGLDREKPEYTLTIQATDMDGSGTTTAVAVVELDANDNAPMFDPOKYEAPHPENA 341
DB 181 IASGLDRVRVETLTWQADLDGQGLTTTALAVIEITVDNDNAPDFDKTYEAAVPE 240
QY 342 VGHEVQRLATVTDAPNSPRAWATYILIMGDDGHDFTTTHPESNQGLTRKGLDFEAK 401
DB 241 AELEVARLATDDEPHTPAWRAVYSIVRNEGAFITTTDPASNEGLRTAKGLDYEAK 300
QY 402 NQHTLYEVTVNEAPFVKLTSTATIVVHVEDVNEAPVFPKSVVEVQEGITGEPCV 461
DB 301 RQFVLHVAVVNEAPFAIKLTATATATVMSVEDVNEAPVFPDPLRLAQVPEDVPLGFLAS 360

QY 462 YTAEDPDK-ENOKISYRILRDPAGWAMDPPDSGGQVTAAGTLDREDEQFVRNNIYVMVLA 520
DB 361 YTAQDPDRAQQRIKYVMGSDPDAGWLAHPENGLITAREQLDRE-SPETKNSTYMAVLA 419
QY 521 WNGSPPTTGTCTLLTLIDVNDHGPVPEPRPOTITCNOSPVHRVLNITDKDLSPHTSFQ 580
DB 420 VDDGLPPATGTGTLTLIDVNDHGPVPEPRDIVICNRSFPQVLTITDRLPPNTGPR 479
QY 581 AQLTDDSDIYTAAYNEEGDTVVLSSKKFLKQDQTDVHLSLSDHGNKEQLTVIRATVDC 640
DB 480 AELSHGSGDSMAVEVNGGDTVALMTEPLEQNLVSVLRLFDROGKQDQTVIRAQVDC 539
QY 641 HGHVETC---POPWIG-GFILLPVLGAVLALLLELLVLLVLRKKRIKEPILLPEDDTRD 696
DB 540 QGRVESCAQKPRVDTGVPVLAVLGAVLALLLLELLVLRKKRIKEPILLPEDDTRD 599
QY 697 NVFYEGEGGGEDDQDITQLHRLGLEARPEVVLVNDVAPTITPTMYRPRPANPDEIGN 756
DB 600 NIFYEGEGGGEDDQDYLUSLHRLGLEARPEVI-RNDVAPPLMAAPQYRPRPANPDEIGN 658
QY 757 FIENLKAAANTDPTAPPYDTLLVFDYEGSGDAASLSLTSSASDQDQDYDLNENWGSRP 816
DB 659 FIDENLKAAADTPTAPPYDLSLVFDYEGGGSEATSLSSLNSSASDQDQDYDLNENWGNRF 718
QY 817 KKLADMYGGEDD 829
DB 719 KKLAEYGGGEDE 731

RESULT 4

IJCCHB

cadherin 1 precursor [validated] - human
N/Alternate names: ARC-1; Cell CAM 120/80; E-cadherin; epithelial cadherin; L-CAM; uvomor
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text change 09-Jul-2004
C/Accession: S37654; S31430; S05475; S31460; S06716; A5171; JC2230; I52294; I52704; S25
R/Bussemaekers, M.J.G.; van Bokhoven, A.; Mees, S.G.M.; Kemler, R.; Schalken, J.A.
Mol. Biol. Rep. 17, 123-128, 1993
A/Title: Molecular cloning and characterization of the human E-cadherin cDNA.
A/Reference number: S37654; MUID:93211394; PMID:8459805
A/Accession: S37654
A/Molecule type: mRNA
A/Residues: 1-882 <BUS>
A/Cross-references: UNIPROT:P12830; UNIPARC:UPI0000341BF; EMBL:Z13009; NID:g31072; PIDN:
R/Kelker, W.; Warda, A.; Oda, T.; Hirohashi, S.; Kemler, R.; Birchmeier, W.
submitted to the EMBL Data Library, December 1992
A/Description: Sequence of human E-cadherin cDNA.
A/Reference number: S31430
A/Accession: S31430
A/Molecule type: mRNA
A/Residues: 1-542, 'F', 544-882 <LE>
A/Cross-references: UNIPARC:UPI000016A86F; EMBL:Z18923; NID:g31074; PIDN:CAA79356.1; PID
R/Mansouri, A.; Spurr, N.; Goodfellow, P.N.; Kemler, R.
Differentiation 38, 67-71, 1998
A/Title: Characterization and chromosomal localization of the gene encoding the human cel
A/Reference number: S05475; MUID:89031725; PMID:3263290
A/Accession: S05475
A/Molecule type: mRNA
A/Residues: 157-311 <MAN>
A/Cross-references: UNIPARC:UPI0000174350; EMBL:X12790
A/Note: nucleotide sequence is not complete
R/Frixen, U.H.
submitted to the EMBL Data Library, March 1990
A/Reference number: S31460
A/Accession: S31460
A/Molecule type: mRNA
A/Residues: 265-392 <PRI>
A/Cross-references: UNIPARC:UPI000016A583; EMBL:X52279; NID:g28821; PIDN:CAA36522.1; PID
R/Wheelock, M.J.; Buck, C.A.; Bechtol, K.B.; Damsky, C.H.
J. Cell. Biochem. 34, 187-202, 1987
A/Title: Soluble 80-kd fragment of cell-CAM 120/80 disrupts cell-cell adhesion.
A/Reference number: S06716; MUID:87280410; PMID:3611200
A/Accession: S06716

A:Molecule type: protein
A:Residues: 'XQ',157-162,'V',164-179 <WHE>
A:Cross-references: UNIPARC:UPI0000174351
R:Berx, G.; Staes, K.; van Hengel, J.; Molemans, F.; Bussemaekers, M.J.G.; van Bokhoven, Genomics 26, 281-289, 1995
A:Title: Cloning and characterization of the human invasion suppressor gene E-cadherin
A:Reference number: A57171; MUID:95324920; PMID:7601454
A:Accession: A57171
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-30,32-882 <BER>
A:Cross-references: UNIPARC:UPI0000174352; GB:L34784
R:Rimm, D.L.; Morrow, J.S.
Biochem. Biophys. Res. Commun. 200, 1754-1761, 1994
A:Title: Molecular cloning of human E-cadherin suggests a novel subdivision of the cadherin
A:Reference number: JC2230; MUID:94242050; PMID:8185635
A:Accession: JC2230
A:Molecule type: mRNA
A:Residues: 1-9,'G',11-15,'RSPILGQSRSPPPCLTRHLVHGAPAPPEKRPR',52-67,'I',69,'LTPPIP',76-94
A:Cross-references: UNIPARC:UPI00001682DE; GB:L08599; NID:G340184; PIDN:AAA61259.1; PID: 1-94
A:Note: the majority of differences between this and other reports represent apparent frameshifts
A:Note: the authors translated the codon CCG for residue 868 as Arg
R:Bussemaekers, M.J.G.; Girolidi, L.A.; van Bokhoven, A.; Schalken, J.A.
Biochem. Biophys. Res. Commun. 203, 1284-1290, 1994
A:Title: Transcriptional regulation of the human E-cadherin gene in human prostate cancer
A:Reference number: 152294; MUID:94380041; PMID:8093045
A:Accession: 152294
A>Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-16 <RES>
A:Cross-references: UNIPARC:UPI000016A870; GB:L34545; NID:G509604; PIDN:AAA21764.1; PID: 1-16
R:Becker, K.F.; Atkinson, M.J.; Reich, U.; Becker, I.; Nekarda, H.; Siewert, J.R.; Hofler, Cancer Res. 54, 3845-3852, 1994
A:Title: E-cadherin gene mutations provide clues to diffuse type gastric carcinomas.
A:Reference number: 152704; MUID:94306394; PMID:8033105
A:Accession: 152704
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 337-476 <RE2>
A:Cross-references: UNIPARC:UPI0000168442; GB:S72492; NID:G632756; PIDN:AAD14108.1; PID: 337-476
A:Comment: Cadherins mediate calcium-dependent intercellular adhesion and are thought to be involved in many cellular processes
C:Genetics:
A:Cross-references: GDB:CDH1; UVO
A:Map position: 16q22.1-16q22.1
A:Introns: 379/3; 440/3
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane protein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:28-154/Domain: propeptide #status predicted <PRO>
F:155-882/Product: E-cadherin #status experimental <MAT>
F:155-697/Domain: extracellular #status predicted <EXT>
F:157-262/Domain: cadherin repeat homology <CR1>
F:232-237/Region: cadherin binding #status predicted
F:265-375/Domain: cadherin repeat homology <CR2>
F:378-486/Domain: cadherin repeat homology <CR3>
F:487-595/Domain: cadherin repeat homology <CR4>
F:596-700/Domain: cadherin repeat homology <CR5>
F:698-731/Domain: transmembrane #status predicted <TM>
F:732-882/Domain: intracellular #status predicted <INT>
F:840-853/Region: serine-rich
F:376,558,570,622,637,849/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 55.4%; Score 2420; DB 1; Length 882;
Best Local Similarity 55.1%; Pred. No. 3.2e-146;
Matches 483; Conservative 112; Mismatches 222; Indels 60; Gaps 13;

QY 9 ASLLQLQV-CWLOCAASPCRAVREAEVTEAGGAEOPGQALGKV-FMGCPG-QEPAL 65
Db 10 ALLLLQLQSSWL-CQEPFPGHGFDAESYTFVPRHLRGLRGLRVNFDCTGRQTAY 68
QY 66 FSTDNDFTVRNGETVQERRSLKERNP----- 92

Db 69 FSLDT-RPKVGTGCVITVRKPLRPHNQIHLVYAWDSTYRKPFSTKVTLNTVGHHRPP 127
QY 93 -----LKI F- SKRILRRHKRDWVAPISVPENKGPFPORLNOLKSNKDRDTK 140
Db 128 HQASVGIQAELETFNNSPGLRQRQDWIPISCPENKGPFPKMLVOIKSNKDEK 187
QY 141 IFYSITPGADSPPEGVAFVEKTEWLLLNKPLDRBEEIAKYELFGHVASNGASVEDPMN 200
Db 188 VFYSITGAGADTFPVGFIERETGMLKVTEPLDRERIAVTTLFHAVSSNGNAVEDPME 247
QY 201 ISIIITDQNDHKPKFTQDTPRGVLEGLPGTSMQVMTATDEDAITYNGVAYSHTSQ 260
Db 248 ILITVTQDNDKPEFTQEVFKGSMGALPGTSMVETATDADDVNTYNAATYTLTSL 307
QY 261 EPKDPHDLMTIHRSTGTSIVISSGLDREKPEVTLTQATDMGDGSGSTTTTAVAVBILD 320
Db 308 DPBLPKMFTINRNTGVSIVVTIGDRESFPYTLVVQAADLQGEGLSTTATAVITVD 367
QY 321 ANDNAPMFDPKQYEAHPENAVGHEVQRLVTVDLADPNSPAWRATYILIMGGDDGHFTIT 380
Db 368 TNDNPFIFNPTTYKGQVPEANVITLKVTDADAPNTFAEAVYTL-NDDGGQFVVT 426
QY 381 THPESNOGILTTAKGLDFEAKNQHTLYVEVTNEAPFVKLPTSTATIVHVEDVNEAPVF 440
Db 427 TNPVNDGILKTAKGLDFEAKQYILHVAVTVVPEVSLTTTATVTVDLVDVNEAPIF 486
QY 441 VPBSKVVEOGIPTGCEPVCVYTAEDPK-ENOKISYRIILRDPAGWLAMPDSCQVTA 499
Db 487 VPPEKRVESDFGVGQEIITSYTAQBPDTMEQRIYIRWRDTANWLEINPDGAI STRA 546
QY 500 TLDREDFQVRNNIYEVMLVMDNGSPPTTGTGTLTLTLDVNDHGVPEVPEPRQITICNQ 559
Db 547 ELDRDFEHVKNSTYTLIATDNGSPVATGTGTLTLLSDVNDNAPIPERTIFFCERN 606
QY 560 PVRHVLNITKDSLPTSPFQAQLTDDSDIYWTAEVNE-BGDTVVLTKLFLKQDTYDVH 618
Db 607 PKQVINIIDALPPNTSPFTAEILTHGASANWITQVNDPTQESIILPKMALEVGDYKIN 666
QY 619 LSLSDHGNKEQLTVIRATVCDCHGVETC--PGPWKGF---ILPVLGAVLALLFLLV 672
Db 667 LKLMNDNQDQVTVLEVSVCDCEGAGVCKAQVPEAGLQIPALGILGILALLILL 726
QY 673 LLLVLRKKRKIKPELLLPEDDTRDNVFFYEGEGGEDDYDITQLHGLEAREPEVVLRN 732
Db 727 LLLFLRRRAVVKPELLPPEDDTRDNVFFYDEEGGEDQDFLSQLHGLDAREVT-RN 785
QY 733 DVAPTIPTMYPRRANPDEIGNFIENKAAINTDPTAPPYDTLLVDFEGSGSDAASL 792
Db 786 DVAPTLMSVPRYLPRRANPDEIGNFIDENLKAADTPTAPPYDLSLLVDFYEGSGSEAASL 845
QY 793 SSLTSSASDQDDYDLNENWGSRFKKLADMYGGGEDD 829
Db 846 SSLNSSESDQDDYDLNENWGNRPFKLADMYGGGEDD 882

RESULT 5
LUMSCE
E-cadherin precursor, epithelial - mouse
N:Alternate names: uvomorulin
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: S04528; S03160; I49565; S48735
R:Nagafuchi, A.; Shirayoshi, Y.; Okazaki, K.; Yasuda, K.; Takeichi, M.
Nature 329, 341-343, 1987
A:Title: Transformation of cell adhesion properties by exogenously introduced E-cadherin
A:Reference number: S04528; MUID:87315445; PMID:3498123
A:Accession: S04528
A:Molecule type: mRNA
A:Residues: 1-412,'V',414-884 <NAG>
A:Cross-references: UNIPROT:P09803; UNIPARC:UPI0000149DC9; EMBL:X06115
R:Ringwald, M.; Schuh, R.; Vestweber, D.; Eistetter, H.; Jottschott, F.; Engel, J.; Doel
EMBO J. 6, 3647-3653, 1987
A:Title: The structure of cell adhesion molecule uvomorulin. Insights into the molecular

A;Reference number: S03160; MUID:98111553; PMID:3501370
A;Accession: S03160
A;Molecule type: mRNA
A;Residues: 157-884 <RIN>
A;Cross-references: UNIPARC:UPI0000174353; EMBL:X06339
A;Note: part of this sequence, including the amino end of the mature protein, was confirmed by R. Behrens, J.; Loewick, O.; Klein-Hitpass, L.; Birchmeier, W.
Proc. Natl. Acad. Sci. U.S.A. 88, 11495-11499, 1991
A;Title: The E-cadherin promoter: Functional analysis of a G-C-rich region and an epithelial enhancer
A;Reference number: I49565; MUID:92107977; PMID:1763063
A;Accession: I49565
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-15 <RES>
A;Cross-references: UNIPARC:UPI000016CC60; GB:M81449; NID:G192325; PIDN:AAA37352.1; PID:R/Tong, K.I.; Yau, P.; Overduin, M.; Bagby, S.; Porumb, T.; Takeichi, M.; Ikura, M.
FEBS Lett. 352, 318-322, 1994
A;Title: Purification and spectroscopic characterization of a recombinant amino-terminal domain of E-cadherin
A;Reference number: S48735; MUID:95010732; PMID:7925993
A;Accession: S48735
A;Status: preliminary
A;Molecule type: protein
A;Residues: 156-300 <TON>
A;Cross-references: UNIPARC:UPI0000174354
C;Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to be involved in many other cellular processes.
C;Genetics:
A;Gene: E-cadherin
C;Superfamily: cadherin; cadherin repeat homology
C;Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane protein
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-156/Domain: propeptide #status predicted <PRO>
F;157-884/Product: E-cadherin, epithelial #status experimental <MAT>
F;157-699/Domain: extracellular #status predicted <EXT>
F;159-264/Domain: cadherin repeat homology <CR1>
F;234-239/Region: cadherin binding #status predicted
F;267-377/Domain: cadherin repeat homology <CR2>
F;380-488/Domain: cadherin repeat homology <CR3>
F;489-597/Domain: cadherin repeat homology <CR4>
F;598-702/Domain: cadherin repeat homology <CR5>
F;702-733/Domain: transmembrane #status predicted <TM>
F;734-884/Domain: intracellular #status predicted <INT>
F;842-855/Region: serine-rich
F;560,639/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 54.7%; Score 2389; DB 1; Length 884;
Best Local Similarity 53.6%; Pred. No. 3.5e-144;
Matches 472; Conservative 123; Mismatches 220; Indels 66; Gaps 13;

Qy 9 ASLLLLQV-CWLOCAASEP--CRAVFREAEVTLAAGAEQEPQALGKV-FMGCPOGEPA 64
Db 10 ALLLLQVSSWL-CQLEPESCSPGFSSEVTFPPVPERHLERGHVLRVRFEGCTGRPT 68
Qy 65 LFSTDNDDFTVRNGETVQERSLK----- 88
Db 69 AFFSEDSRFKATDGTITVTKRHLKHLKLETSFLVRARDSSHRELSTKVTLSKNGHHHRH 128
Qy 89 -----ERNP-LKIFPPSKRI-LRRHKRDWVAPISVPENGKGFPPORLNQKSNKDRDT 139
Db 129 HHRDPASENPPELLMFPSPVGLRRQKRDWVIPPSCPENKEGEPFKQLVQIKSNDRKET 188
Qy 140 KIFYSITGQADSPGCVFAVEKGTGMLLNKPLDREIAKTELFGHVAENGASVEDPM 199
Db 189 KVFYSITGQADKPPVGVFIETRTGWLKVTQPLDREAIAKTYLSHAVSNSGEAVEDPM 248
Qy 200 NISIVTDQNDHKPKFTQDFRGSLVGLVPGTSMQVATDDEDAITYNGVAVSIHS 259
Db 249 EIVITVTDQNDNRPEFTQEVFEGSAEAGVPGTSVMKVSATDADDVNTYNAIAIYVS 308
Qy 260 QEPKDPHDLMTTHRSVTGTTISVSSGLDREKVPETLTITQATDMDGSGTTFVAVVEIL 319
Db 309 QDELPHPKMWFTVNRDGTGLSVLTSLGLDRESYFTYTLVQVQADLQEGSLSTAKAVITVK 368
Qy 320 DANDNAPMFQKYEAHVPENAVGHEVQRLTVTDLPNPSPAWRATYLIIMGDDGDGHFTI 379

Db 369 DINDNAPFNPSTYQGVPENEVNARIATLTKVTDADPNTPAWKAVYTVV-NDPQQQFVV 427
Qy 380 TTHPESNOGILTRKGLDFEAKNQHTLYVEVNEAPFVLKLPSTSTATIVVHVVEDVNEAPV 439
Db 428 VTDPITNDGILKTAGLDFEAKQVILHVRVNEEFEGSLVPSTATVTVDDVVDNEAPI 487
Qy 440 FVPPSKVVEVQGIPTGPEVCVYADBDPK-ENQKISVRIILDPAGWLAMDPSGQVAV 498
Db 488 FMPAERRVEVPBDFGVQGBITSYTAREPDTFMDQKITVIRMRDRTANWLEINPETAIFTR 547
Qy 499 GTLDREDEQFVNNIYEVNVLAMDNGSPPTCTGTLLTLTLDVNDHGPVPERQITICNQ 558
Db 548 AEMDRDEAHVKNSTVALIITDDGSPATGTGILLVLLDVNDONAPIPERNMQFCOR 607
Qy 559 SPVRHVNLNTKDLSPHTSPFOQLTDDSDIYWTAEVNEEG-DTVVLSLKFKLKQDTYDV 617
Db 608 NPQPHIITLDPDLPNTSPFTAEHLTHGASVNMVTEYNDAAQESLILQPKLEIGEYKI 667
Qy 618 HUSLDHGNKEQLTVIRATVCDCHGHVETCPGPWKGFF-----ILPVLGAVLALLF 668
Db 668 HLKLDNQNKQVTTLDVHVCDCEGTNNC---MKAGIVAAGLQVPAIILGILGILALLI 724
Qy 669 LLLVLLLVKRRKIKPELLPEDDTRDNVYFYEGEGGEDDQDITOLHRLGLARPEV 728
Db 725 LILLLLFURRTVVKPEPLPPDDTRDNVYFYDEEGEGGEDDQDPLSQHLRLGLARPEV 784
Qy 729 VLNRNVAPTIITPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVFDYEGSGD 788
Db 785 T-RNDVAPFLMSVQVPRPANDPDEIGNFIDENLKAADSPTAPPYDLSLVFDYEGSGSE 843
Qy 789 AASLSLSTSSAQDQDQDYDLNWSRPFKKLADMYGGGEDD 829
Db 844 AASLSLSTSSAQDQDQDYDLNWSRPFKKLADMYGGGEDD 884

RESULT 6
S34438
uvomorulin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Nov-1993 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S34438
R;Ringwald, M.; Baribault, H.; Schmidt, C.; Kemler, R.
Nucleic Acids Res. 19, 6533-6539, 1991
A;Title: The structure of the gene coding for the mouse cell adhesion molecule uvomorulin
A;Reference number: S34438; MUID:92093614; PMID:1754391
A;Accession: S34438
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-884 <RIN>
A;Cross-references: UNIPROT:P09803; UNIPARC:UPI0000020A82; EMBL:X60975
A;Note: the nucleotide sequence was submitted to the EMBL data library, August 1992
C;Genetics:
A;Intons: 16/3; 57/1; 131/3; 179/3; 231/3; 280/1; 338/3; 381/3; 442/3; 524/2; 573/1; 646/1
C;Superfamily: cadherin; cadherin repeat homology
F;380-488/Domain: cadherin repeat homology <CR3>

Query Match 54.5%; Score 2379; DB 2; Length 884;
Best Local Similarity 53.5%; Pred. No. 1.3e-143;
Matches 471; Conservative 122; Mismatches 222; Indels 66; Gaps 13;

Qy 9 ASLLLLQV-CWLOCAASEP--CRAVFREAEVTLAAGAEQEPQALGKV-FMGCPOGEPA 64
Db 10 ALLLLQVSSWL-CQLEPESCSPGFSSEVTFPPVPERHLERGHVLRVRFEGCTGRPT 68
Qy 65 LFSTDNDDFTVRNGETVQERSLK----- 88
Db 69 AFFSEDSRFKATDGTITVTKRHLKHLKLETSFLVRARDSSHRELSTKVTLSKNGHHHRH 128
Qy 89 -----ERNP-LKIFPPSKRI-LRRHKRDWVAPISVPENGKGFPPORLNQKSNKDRDT 139
Db 129 HHRDPASENPPELLMFPSPVGLRRQKRDWVIPPSCPENKEGEPFKQLVQIKSNDRKET 188

```
Qy 140 KIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDREBIEAKYELFGHAVSENGASVEDPM 199
Db 189 KVFYSITGQADKPPVGVFIIRERETGMLKVQPLDREALAKIYLYSHAVSSNGEAVEDEPM 248
Qy 200 NISIIVTDQNDHKPKFTQDTPRGSLVGLPDTGSMQVMTATDEDAIYTYNGVVAIHS 259
Db 249 EIVITVTQDNDNRBEFTQPVPEGFVAEGAVPGTSMVKVSAFTDADDVNTYNAAIAYTIVS 308
Qy 260 QEPKDPDLMTIHRSTGTSIVISGLDREKVPYVTLTIQATMDGSGSTTAVAVVEIL 319
Db 309 QDPELPHKNMFTVARDTGIVSLVSGLDRESYPYTLVVQAAQGGSLSTAKAVITVK 368
Qy 320 DANDNAPFDPQKYEAHPENAVGHEVORLVTVTDLDAPNSPAMWATYLIIMGDDGDHFTI 379
Db 369 DINDNAPVFNSTYQGVQVPEVNAVARATLKVTDADAPNTPAWKAIVTVV-NDPDQGFVV 427
Qy 380 TTHPESNOGILTTAKGLDPEAKNOHTLYVEVTNAPVFLKLPSTTATIVHVEDVNBAPV 439
Db 428 VTDPTTNDGILKATKGLDPEAKQYILHVRVENEPEFEGSLVPSTATVTVDDVDVNEAPI 487
Qy 440 FVPESKVEVOEGIPTGEPVCVYTAEDPK-ENOKISYRILRDPAGWLAMPDPSGOVTAV 498
Db 488 FMPAERRVEVPEDFGVGOEITSYTAREPDTTMDQKITRYIWRDTANLWLEINPETGAIFTR 547
Qy 499 GTLDREDEQFVRNNIYEVWVLAMONGSPPTTGTGTLTLIDVNDHGFVPPEPRQITICNQ 558
Db 548 AEMDREDAEHVKNSTYVALIATDGGSPATGTGTLVLVLDVNDNAPIPEPRNMQFCQR 607
Qy 559 SPVRHLNITDKLSPTSPFOALTDSDSIYVTAENVEEG-DIVVLSLKKFLKQDYDV 617
Db 608 NPQPHIITLDPDLPNTPSPTELTGASVNWITIEYNDAAQESLILQPKRDLGEIGYKI 667
Qy 618 HLSLSHCNKQLTVIRATVCDCHVETCPGPKKGF-----ILPVLGAVTALLF 668
Db 668 HKLADNQNQDQVTLTVHVCDCGTVNNC---MKAGIVAAAGLVPAILGLGILALLI 724
Qy 669 LLLVLLLVKRRKIKPELLLPEDDTRDNVYVYEGEGGEDQDYDTQLHRGLEARPEV 728
Db 725 LILLLLFLRRRTVVKPELLPDDTRDNVYVYDEEGGEDQDFDLSQLHRGLDARPEV 784
Qy 729 VLRNDVAPTIPTMYRPRRANPEIGNFIENKAAWTDPTAPYDTLLVDFYEGSGSD 788
Db 785 T-RNDVAPTLMSVQYRPRRANPEIGNFIENKAAWTDPTAPYDLSLVDFYEGSGSE 843
Qy 789 AASLSLTSASDQDQDYDLNENSGSRKKLADMYGGEDD 829
Db 844 AASLSLNSSESDDQDQDYDLNENGNRPFKLADMYGGEDD 884

RESULT 7
I3XLCP
EP-cadherin precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: B43785; A60128
R:Ginsberg, D.; Desimone, D.; Geiger, B.
Development 111, 315-325, 1991
A:Title: Expression of a novel cadherin (EP-cadherin) in unfertilized eggs and early xenopus oocytes
A:Reference number: A43785; PMID:91372132; PMID:1893866
A:Accession: B43785
A:Molecule type: mRNA
A:Residues: 1-895 <GIN>
A:Cross-references: UNIPROT:P33148; UNIPARC:UPI0000171531; GB:X63720; NID:g64681; PIDN:C
A>Note: it is uncertain whether Met-1 or Met-16 is the initiator
R:Angres, B.; Mueller, A.H.J.; Kellermann, J.; Hausen, P.
Development 111, 829-844, 1991
A:Title: Differential expression of two cadherins in Xenopus laevis.
A:Reference number: A60128; PMID:91347911; PMID:1879345
A:Accession: A60128
A:Molecule type: protein
A:Residues: 171-177, 'I', 179-183, 'K', 185-189, 'XI' <ANG>
A:Cross-references: UNIPARC:UPI000017435C
A>Note: the material sequenced may have contained U-cadherin as well as E-cadherin
```

```
C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; embryo; glycoprotein; transmembr
F:1-43/Domain: signal sequence #status predicted <SIG>
F:44-170/Domain: propeptide #status predicted <PRO>
F:171-895/Product: EP-cadherin #status predicted <MAT>
F:171-718/Domain: extracellular #status predicted <EXT>
F:173-278/Domain: cadherin repeat homology <CR1>
F:248-253/Region: cadherin binding #status predicted
F:281-391/Domain: cadherin repeat homology <CR2>
F:503-610/Domain: cadherin repeat homology <CR3>
F:611-714/Domain: cadherin repeat homology <CR4>
F:719-743/Domain: transmembrane #status predicted <TM>
F:744-895/Domain: intracellular #status predicted <INT>
F:852-865/Region: serine-rich
F:440,696/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 54.4%; Score 2375; DB 1; Length 895;
Best Local Similarity 54.2%; Pred. No. 2,4e-143;
Matches 471; Conservative 128; Mismatches 214; Indels 56; Gaps 9;

Qy 11 LLLQVWCWQCAASEPCRAVFEAEVTLGAGAEQEPQALGV-FMCGCPQEPALPSTD 69
Db 32 LCLLVVPSINADVSGCKPGFSSAEYIFSVNRRELERGRKLGKVFSDCTTRKHGLYVG 91
Qy 70 NDDFTV-----RNGETVQERRS 86
Db 92 DSRFVLPGDTVLVKRHVKLHKDTFTISTWDARGIKHSTNIAVAKRHRSGEAIHRSR 151
Qy 87 LKERNLKIPPSKRI-LRRHKRDWVAPISVPENKGGPFPQRLNQLKSNKDRDTKIYSI 145
Db 152 ---KLPLVLTPEHTGLKRRKRDWIPPIKVSENERGFPKRLVQIKSNKDRNKVYSI 208
Qy 146 TGGASDPGPGVFAVKETGWLNLNKLPLDREBTAKYELFGHAVSENGASVEDPMNISII 205
Db 209 TGGADNPQGVRIEWETGWLTVRPLDREEDYKYVLSHAVSENGSPVEEPMETINV 268
Qy 206 TDQNDHKPKFTQDTPRGSLVGLPDTGSMQVMTATDEDAIYTYNGVVAIHSOEPKP 265
Db 269 IDQNDNRPKFTQDTPRGSLVGLPDTGSMQVMTATDEDAIYTYNGVVAIHSOEPKP 328
Qy 266 HDLMFTIHRSTGTSIVISGLDREKVPYVTLTIQATMDGSGSTTAVAVVEILDANDNA 325
Db 329 IPNLFTINRTGTVISLIGTGLDREKPEYTLTVQATDLEGAGLSVEGKAIITDANDNA 388
Qy 326 PMFDPPQKYEAHPENAVGHEVORLVTVTDLDAPNSPAMWATYLIIMGDDGDHFTIHPES 385
Db 389 PIFDPKTYTALVPENEIGFEVQRLSVTDLMPGTPAMQAVYKIR-VNEGGFNITTDPS 447
Qy 386 NOGILTRKGLDPEAKNOHTLYVEVTNAPVFLKLPSTTATIVHVEDVNEAPVFPVPSK 445
Db 448 NOGILTRKGLDPEAKNOHTLYVEVTNAPVFLKLPSTTATIVHVEDVNEAPVFPVPSK 507
Qy 446 VVEQEGIPTGEPVCVYTAEDPKEN-QKISYRILRDPAGWLAMPDPSGOVTAVGTLDRE 504
Db 508 RVDVSEDLRGEKISILVAQDPKQIQKLSYFIGNDPARWLTVNKDNGIVTGNGLDRE 567
Qy 505 DEQFVRNNIYEVWVLAMONGSPPTTGTGTLTLIDVNDHGFVPPEPRQITICNQSPVRHV 564
Db 568 SE-YVKNITYTVIMLVTDGVSVGTGTGLIHLVLDVNDNGVPSPRVFTMCDQNPQV 626
Qy 565 LNIITDKLSPHTSPFOALTDSDIYVTAENVEEGDTVLVLKFLKQDYTVHLSLSDH 624
Db 627 LTISDADIPNTTYPYKVSLSHSGDLTWKAEKSGTSMLLSPQQQLKGGYIYVLUSDA 686
Qy 625 GNKEQLTVIRATVCDCHVETCPGPKKGFILP----VLGAVLALLFLLLVLLLVXX 680
Db 687 QNNPQLTVVNAVTCSCGKAIKQEKLVGGFDLPILVILGSLVALLILFLLLLFLKRX 746
Qy 681 RKIKEPILLPDDTRDNVYVYEGEGGEDQDYDTQLHRGLEARPEVLRNDVAPTIIP 740
Db 747 KVKKEPILLPDDTRDNVYVYEGEGGEDQDYDLSQLHGLDSRPD-IMRNDVVPITMP 805
```

```
QY 741 TMYRPRPNDPDEIGNFIENLKAANTDPTAPPYDTLLVPDYEGSGDAASLSSLTSSAS 800
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 806 APHYRPRSPNDPDEIGNFIDENLDAANDPTAPPYDSILLVPDYEGSGEAAASLSSLNSSNS 865
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 801 DQDQDYDLNEWGSRPKKLADMYGGGDD 829
      : : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 866 NDEHDNYLSDWGSFRKRLADMYGGDDDE 894
      : : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
S43064
cadherin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Aug-1999
C:Accession: S43064; S43065
R:Kuehl, M.
submitted to the EMBL Data Library, March 1994
A:Reference number: S43064
A:Accession: S43064
A:Molecule type: mRNA
A:Residues: 1-905 <XUE>
A:Cross-references: UNIPARC:UPI00001715C5; EMBL:X78546; NID:G468816; PIDN:CAA55292.1; PI
R:Harzberg, F.; Wildermuth, V.; Wedlich, D.
Mech. Dev. 35, 33-42, 1991
A:Title: Expression of XBCad, a novel cadherin, during oogenesis and early development
A:Reference number: S43065; MUID:92062581; PMID:1840622
A:Accession: S43065
A:Molecule type: mRNA
A:Residues: 'NSA', 462-697, 'Q', 699-807, 'A', 809-840, 'N', 842-877, 'N', 879-883, 'N', 885-902, 'E'
A:Cross-references: UNIPARC:UPI0000171511; EMBL:X78546
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; glycoprotein; transmembrane protein
F:181-286/Domain: cadherin repeat homology <CR1>
F:289-399/Domain: cadherin repeat homology <CR2>
F:402-510/Domain: cadherin repeat homology <CR3>
F:511-618/Domain: cadherin repeat homology <CR4>
F:619-722/Domain: cadherin repeat homology <CR5>

Query Match 54.3%; Score 2371; DB 2; Length 905;
Best Local Similarity 54.1%; Pred. No. 4.4e-143;
Matches 469; Conservative 125; Mismatches 222; Indels 51; Gaps 8;

QY 11 LLLQVLCWLOCAASEPCRAVFREAEVTLLEAGAEQEPGQALGV-FMGCGQBPALFSTD 69
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 39 LCLQVVPISNDVSGCQPGFSSANVTFSVNRLELGRKLGKVLVDCVTRKHGLYVG 98
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 70 NDDFTVRNGTV-----QERRSLK 88
      : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 99 DSRFRVLPGDTVLVGRVHLKSKDTRFTISTWDARGIKHSTNISVYVNRHRSGEARSRS 158
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 89 ERNPLKIFPSKRI-LRRHKEDWVAPISVPENCKGPPQPPORLNOLKSNKORDTKIFYSITG 147
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 159 SELPVLVTFPEKHGLKRRKRDWIPPIKVSNERBERGPPFKRLVQIKSKKELSKVFSYITG 218
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 148 PGADSPPEGVAFVEXETGWLNLNKLNDREIAKELFCHAVSNGASVEDPMNISIVTD 207
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 219 QGADTPPEGIFREKETGHWQVTRPLDREYKYLSSHAVSNGASVEEPMEITVVID 278
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 208 QNDHKPKFTQDTRFGSVLEGLVPGTSVMQVATDEDAIYTVNGVAYSIHSGEPKDPHD 267
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 279 QDNDRPKFTQVPRFSGVREGVPGTKVMSVATDDSDSLNGLVAYSILKODPEEPIP 338
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 268 LMFTHIRSTGTISVISGLDREKVPEYTLTIQATDMDGSGSTTAVAVVEILLDANDNPM 327
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 339 NLFTINRETGVISLIGLDRKPEPPEYTLVQAADLQAGLTAEGKAVIEITDANDNAPI 398
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 328 FDPQKEAHYPENAVGHEVORLVTDLDAFNSPAWRATYLLIMGGDGDHFTITTHPESNQ 387
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 399 FDPKTYTALVPENVEGFEVORLSVTDLMDPGTAQWQVYKIR-VNEGGFNITDPESNQ 457
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 388 GILTTAKGLDPEAKNOHTLVVEVNEAPFVLKLPSTATVHVHVEDVNEAPVFPVPSKV 447
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 458 GILTTAKGLDPEVRKQYVLIQITVENAVPSPVPLPTSTATVTVTVEDVNEAPVFPVPSRV 517
QY 448 EYQEGTPTGEPVCVYTAEDPKEN-QKISYRILRDPAGWLAMPDPSGQVTVAGTLDREDE 506
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 518 DVSLEDLRGEKIVSLVAQDPDKQIQKLSYFIGNDPARWLTKINGIVTGNGLDRESE 577
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 507 QVFRNNIYEVWVLAAMDNGSPPTGTGTLTLLIDVNDHGPVPEPRQITTCNSQVPRHVLN 566
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 578 -YVKNTYTVIMLVTDGVPVGTGTLLHLVLDINDNGPVPSPRVFTWCDQNPQPVLT 636
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 567 ITDKDLSPTSPFQAQLTDDSDIYKTAEVNEEGDVTVLSSKKFLKODTVDVHLSLDHGN 626
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 637 ITDADIPNTYPSVLSHGSELTKWAEUSDGKTSWRLSPQTQQLKKGDSIYIVLLADAQA 696
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 627 KQQLTVIRATVCDCHGVETCFPGPWKGGFILP-----VLGAVLALLFLLLVLLLVKRRK 682
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 697 NRQLTVNATVCIPEGKAIKCEKLVAGFDLPILVILGSILALLLSLLLLFLKRRKV 756
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 683 IKPEPLLPEDDTRDNVFFYGGEGGEEDQDYITQLHRLGEARPEVVLVNDVAPTIIP 742
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 757 VKEPLLPEDDTRDNIFYYGEGGGEEDQDYDLSQLHRLGDARPD-IMENDVVPVTLMSVP 815
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 743 MYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVPDYEGSGDAASLSSLTSSASQ 802
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 816 HVRPRSPNDEIGNFIDENLDAANDPTAPPYDSLLVPDYEGSGEAAASLSSLNSSNN 875
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 803 DQDYDLNEWGSRFKKLADMYGGGDD 829
      : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 876 EHDYNYLNDMGPRFKLADMYGGDDDD 902
      : |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
IJBQCP
P-cadherin - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: S11694
R:Liaw, C.W.; Cannon, C.; Power, M.D.; Kiboneka, P.K.; Rubin, L.L.
EMBO J. 9, 2701-2708, 1990
A:Title: Identification and cloning of two species of cadherins in bovine endothelial cel
A:Reference number: S11693; MUID:90360979; PMID:2390969
A:Accession: S11694
A:Molecule type: mRNA
A:Residues: 1-491 <SLIA>
A:Cross-references: UNIPROT:P19535; UNIPARC:UPI0000126D92; EMBL:X53614; NID:G166; PIDN:CI
C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane prot
F:1-491/Product: P-cadherin (fragment) #status predicted <EXT>
F:1-314/Domain: extracellular (fragment) #status predicted <EXT>
F:1-102/Domain: cadherin repeat homology (fragment) <CR3>
F:103-210/Domain: cadherin repeat homology <CR4>
F:211-314/Domain: cadherin repeat homology <CR5>
F:315-339/Domain: transmembrane #status predicted <TMM>
F:340-491/Domain: intracellular #status predicted <INT>
F:447-462/Region: serine-rich
F:228/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 53.2%; Score 2323; DB 1; Length 491;
Best Local Similarity 88.4%; Pred. No. 2e-140;
Matches 434; Conservative 23; Mismatches 34; Indels 0; Gaps 0;

QY 339 ENAVGHEVORLVTDLDAFNSPAWRATYLLIMGGDGDHFTITTHPESNOGILTRKGLDF 398
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 ENTVSHEVORLVTDLDAFNSPAWRATYRIVGDNQGDHFTITTHPESNOGILTRKGLDF 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 399 EAKNOHTLVVEVNEAPFVLKLPSTATVHVHVEDVNEAPVFPVPSKVVEGIGTGP 458
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 EAKTQHTLVVEVINEVPFVKLPSTATVVLVVEDVNEPVPFVPSKVIEIGISTGP 120
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 459 VCVYTAEDDPKQKISYRILRDPAGWLAMPDPSGQVTVAGTLDREDEQVFRNNIYEVW 518
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 ICAYTARDPKGSQKISYHILRDPAGWLAMPDPSGQVTAAGVLDREDEQVFRNNIYEVW 180
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Qy 519 LAMNGSPPTTGTCTLLTLTLDVNDHGVPEPRQITTCNOSPVRHVLNITDKDLSPTSP 578
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 LATDGGSPPTTGTCTLLTLTMDNDHGVPEPRQITTCNOSPVPQVNLITDKDLSPTAP 240

Qy 579 FQALTDSDSIYTAEVNEEGDVTYLSLKKFLKQDITYVHLSLSDHGKKEQLTVIRATVC 638
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 FQALTHDSDVYTAEVNEKGDVAVALSLKKFLKQGEYDVHLSLSDHGKKEQLTVIRATVC 300

Qy 639 DCHGHVETCPGWKGFIPLVGLAVLALLFLVLLVLLVLRKKIKEPFLLPEDDTRDNV 698
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 DCHGNMVTCDPWTWGFLLPTLGLAALALLLLVLLVLRKKIKEPFLLPEDDTRDNV 360

Qy 699 FYYEGGGEDDQDYDITQLHRLGLEARPEVVLNDVAPTITPTMYPRPANPDEIGNFI 758
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 FYYEGGGEDDQDYDITQLHRLGLEARPEVVLNDVAPSFIPTMYPRPANPDEIGNFI 420

Qy 759 IENLKAANTDPTAPPYDTLLVDFYEGSGSDAASLSSTSSASDQDQDYDYNLWNGSRFKK 818
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 IENLKAANTDPTAPPYDLSLLVDFYEGSGSDAASLSSTSSDQDQDYDYNLWNGSRFKK 480

Qy 819 LADMYGGGEDD 829
|||||:|||||:
Db 481 LADMYGGGQDD 491

RESULT 10
LUCGCL
E-cadherin precursor, hepatic - chicken
N:Alternate names: L-CAM; liver cell adhesion molecule
C:Species: Gallus gallus (chicken)
C:Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: A30201; A29866; E99866
R:Sorkin, B.C.; Hemperly, J.J.; Edelman, G.M.; Cunningham, B.A.
Proc. Natl. Acad. Sci. U.S.A. 85, 7617-7621, 1988
A:Title: Structure of the gene for the liver cell adhesion molecule, L-CAM.
A:Reference number: A30201; MUID:89017248; PMID:3174655
A:Accession: A30201
A:Molecule type: mRNA
A:Residues: 1-81 <SOR>
A:Cross-references: UNIPROT:P08641; UNIPARC:UPI0000174355; EMBL:J04074
R:Gallin, W.J.; Sorkin, B.C.; Edelman, G.M.; Cunningham, B.A.
Proc. Natl. Acad. Sci. U.S.A. 84, 2808-2812, 1987
A:Title: Sequence analysis of a cDNA clone encoding the liver cell adhesion molecule, L-
A:Reference number: A29866; MUID:87204217; PMID:3472238
A:Accession: A29866
A:Molecule type: mRNA
A:Residues: 51-887 <GAL>
A:Cross-references: UNIPARC:UPI0000174356; EMBL:M16260
A:Accession: B29866
A:Molecule type: protein
A:Residues: 161-172;323-336;386-407;533-551 <GA2>
A:Cross-references: UNIPARC:UPI0000174357; UNIPARC:UPI0000174358; UNIPARC:UPI0000174359;
C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
C:Genetics:
A:Introns: 23/3; 62/1; 138/1 183/3; 235/3; 284/1; 342/3; 385/3; 447/3; 529/2; 577/1; 652
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; liver; transmembr
F:1-26/Domain: signal sequence #status predicted <SIG>
F:161-160/Domain: propeptide #status predicted <PRO>
F:161-887/Product: E-cadherin, hepatic #status experimental <MAT>
F:161-704/Domain: extracellular #status predicted <EXT>
F:163-268/Domain: cadherin repeat homology <CR1>
F:238-243/Region: cadherin binding #status predicted
F:271-381/Domain: cadherin repeat homology <CR2>
F:384-493/Domain: cadherin repeat homology <CR3>
F:494-601/Domain: cadherin repeat homology <CR4>
F:602-704/Domain: cadherin repeat homology <CR5>
F:705-735/Domain: transmembrane #status predicted <TM>
F:736-887/Domain: intracellular #status predicted <INT>
F:291,346,564,643/Binding site: carbohydrate (Asn) (covalent) #status experimental
```

```
Query Match 52.0%; Score 2274; DB 1; Length 887;
Best Local Similarity 52.7%; Pred. No. 6-6e-137;
Matches 461; Conservative 122; Mismatches 236; Indels 55; Gaps 13;

Qy 8 LASLLLVQCWLOCAASEPCRAVREAEVTLTAGGAEQEPQALGKV-FMGCGQGPALF 66
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 16 LVTLVLLVCGRRCDAAACQPGFAAETFSFSPQDVAAGRELGRVSAACSGRPWAVY 75

Qy 67 -STD-----NDP-----FTVRNGETVQERRS----- 86
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 76 VPTDTRFKVNGDGVSTKRPPLTYLGRKKISFTIYAQAMGKRHSARVTVGRHRRHHNH 135

Qy 87 -LKERNPLKI-FP--SKRILRRHRDWMVAPISVPENKGGPPQRLNQLKSNKDRDTKIF 142
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136 HLQDTPPAVLTFPKHPDGLRRQKRDWVIPPISCLNHRGPYPMRLVQIKSNKDKSKVY 195

Qy 143 YSTGCGASDPPEGVAFAVETGWLINLKNPLDREEIAKYELFGHVAHSENGASVEDPMNIS 202
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 196 YSITGCGADSPPGVIGFIIEBRETGWLEVTQGLDRKIDRYTLLSHAVSASGQVEDPMELI 255

Qy 203 IIVTDQNDHKPKFTQDTERGSLVLEGLPGTSVMQVATDDEDAIYTYNGVVAYSIHSQEP 262
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 256 ITVMDQNDKPKVFIEKVFVGYIEENAKPGTSVMVNTADDAVNTDNGIVSIVSQQP 315

Qy 263 KDPHDLMTIHRSTGTISVISGLDREKVPETLTITQATDMGDGSTTTAVAVVEILDAN 322
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 316 PRHPQWFTIDPAKGIIISVLGTGLDRETTNPYTLIVQATDQEGKGLSNTAIIIEVDAN 375

Qy 323 DNAPMFDPOKYEAHVPENAVGHVQRLTVTDLDAPNSPAWRATYILMGDDGDGHFTITTH 382
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 376 DNIPIFNPTMYEGWVENKPGTEVARLTVTDQAPGSPAWQAVYHIKSGNLGDAFSIITD 435

Qy 383 PESNOGILTKRGLDFEAKNQHTLYVEVTNEAPFVLKLPSTSTATIVVHVEDVNEAPVFP 442
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 436 PSTNNGILTKAKGLDYETKSDLVVTVENKVPSPITLSTASVLTVLDVNEPPVFP 495

Qy 443 PSKXVEVEQIGTGPVVCVTAEDDPKE-NQKISYRILRDPAGLWMDPDGQVAVGTIL 501
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 496 PIKRVGVPEDLPVGOQVTSYTAQDPDRDMRQKITRMSGSDPAGWLYIHPENGIVTATQPL 555

Qy 502 DREDEQFVRNIIYEVWVLMNDGSPPTTGTGTLTLLTLDVNDHGVPEPRQITICNQSPV 561
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 556 DRESVHAI-NSTYKAILAVDNGIPDTGTGTLTLLQLDQVNDNGPTPEPRSFICSRQPE 614

Qy 562 RHYLNITDKDLSPTSPQALTDSDIYVTAENVEGDTVLSLKKFLKQDYVDVHLSL 621
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 615 KQILSIVDKDLPPTPTPFKAALHSGSNWTVIRGQ-DELAAGLKKLELPGEYNIFVKL 673

Qy 622 SDHGKNEQLTVIRATVCDCHGVETCP--GPMWKGGF-----ILPVLGAVLALLLVL 675
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 674 TDSQGAQVTVQKVAQVCECEGTAKNCERRSYIVGGLGVPAIILGILGAILALLL 733

Qy 676 LVKKRKIKEPFLLPEDDTRDNVYFYGEGGGEDDQDYITQLHRLGLEARPEVVLNDVA 735
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 734 FARRRKVEKEPLLPEDDMDRDNVYNYDEEGGGEDQDYDLSQLHRLGLDARPEVI-RNDVA 792

Qy 736 PTIIPMPYPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVDEYEGSGSDAASLSL 795
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 793 PPLMAAPQYPRPANPDEIGNFIDENLKAANTDPTAPPYDLSLLVDFYEGGGSEATSLSL 852

Qy 796 TSSASDQDQDYDYNLWNGSRFKKLADMYGGGEDD 829
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 853 NSSASDQDQDYDYNLWNGNRFKKLAEIYGGGEDD 886

RESULT 11
S47518
cadherin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 01-Feb-1995 #sequence revision 10-Nov-1995 #text_change 17-Mar-1999
C:Accession: S47518
R:Tooi, O.; Fujii, G.; Tashiro, K.; Shikawa, K.
Biochim. Biophys. Acta 1219, 121-128, 1994
```

A;Title: Molecular cloning of cDNA for XTCAD-1, a novel Xenopus cadherin, and its expres
A;Reference number: S47518; MUID:94368839; PMID:8086449
A;Accession: S47518
A;Molecule type: mRNA
A;Residues: 1-871 <TOO>
A;Cross-references: UNIPARC:UPI0000177AF8
C;Superfamily: cadherin; cadherin repeat homology
F;151-256/Domain: cadherin repeat homology <CR1>

Query Match 50.2%; Score 2192.5; DB 2; Length 871;
Best Local Similarity 51.9%; Pred. No. 1e-131;
Matches 460; Conservative 126; Mismatches 227; Indels 73; Gaps 16;
QY 1 MGLPR-----GPLASLLLLQVC-----WLQCAASEPCRAVFEAEVLEAGAEQPGQAL 51
DB 1 MGLKRPWLLGAVVLLTLIQVGLAEWTQ-----CRNGFSKEKYSFLV-PKNLETDKAL 53
QY 52 GKV-FWGCPCQEPALSTDNDDFTVRNGETVQRRSLKERNPLKIP----- 96
DB 54 GRVFNCEGVRVRIQFASKDPNFEIHKDGTYYIKNPAKMDNKRKTRVLAWETKGHVYST 113
QY 97 -----PSKRILRRHKRDWVAVIPGVPENGKGPFPORLNOLK 132
DB 114 NITLKREHHRQDLFSGKSHHPKSETGLKQKRDWVIPPVIVSENEKGPFPKRIVQIK 173
QY 133 SNKDRDTKIFYSTGPGADSPPEGVFAVEKETGMLLNKPLDREETAKYELFGHVSNG 192
DB 174 SSVAKEVKVYISITGQADTPPEGVFAIGREDGWLNVTRPLDREAINDVYLFPSHAVSSNG 233
QY 193 ASVEDPMNISIIITDQNDHKPKTKTQDTPRGSVLGVLPGTISVMQVATDEDDAIYTYNGV 252
DB 234 ANYEDPMETIIKVDQNDNDPFTVQTSFEGSVPEGSKPGTAVTVSATDADSDVMYNGV 293
QY 253 VAYSISQEPKPDHLMFTHTSTGTSISVSSGLDREKPEYTLTIQATMD--GDGSTTT 311
DB 294 ITSILNQEPKEPTNKRFTIHSSEGLSVLTGLDREKNPVYTLTQAADGEGFKDRTTT 353
QY 312 AVAVVEILDANDNAPDPPOKYEAHVPENAVGHEVQRLTVTDLDAPNSPAWRATYILMG 371
DB 354 ATALIVUMDTNDNPPVDPPTQYAKVPENEVGEVARTVTDEIDEGTDAMNAVYKIIG 413
QY 372 DGDGHTFITTHPSNGILTRKGLDPEAKNQHTLVYEVTNEAPFVLKLPSTATTIVHV 431
DB 414 NEANYESIQT-DTGNIGLLKTVKGLDYELKQVILSVIVTNKANFSVPLQTSTATVTSV 472
QY 432 EDVNEAPVFPSPKVEVQSGIPTGEPVCVYTAEDPDKE--NOKISYRILRDPAGWLMDP 490
DB 473 EDVNEAPFIFLPKVEVSSEDLSFGQVATYTTAQDPDKQKQKITTVIGNDPSRWVSVNK 532
QY 491 DSGQVAVGLDREDEQFVRNNITYEVMVLAMDNQSPPTTGTGTLTLTLDVNDHGPVPEP 550
DB 533 DNGIVTGNGLDRE-SKFLVNTYKVIILAADSGSPSATGTGTLVNLNLDVNDNGFFLEP 591
QY 551 RQITCNQSPVRHVLITDKDLSPTSPFOALTDSDIYWTAEVNEEGTVVLSL--KK 608
DB 592 QBSFQKQDQGFVFTIIDRLSPNTYKAEITGESNENWTAIVTDK----ILELRPKK 647
QY 609 FLAQDTVDVHLSISDRGNKEQLTVIRATVCDCHGVETC--PGPWKGF-----ILVLGA 662
DB 648 ELEIGQVDMITLDSFGVSNVTKLHITICQCDGDKKQCEKAIAAGGLGISAIVGILGG 707
QY 663 VLALLFLLVLLLVKRRKIKPEPLLPEDDTRDNVPYFEEGGEGEDQDYDTQLHRGL 722
DB 708 ILALLLLLLLLLVKRVKVEPLPPEDETRDNVFSYDEEGGEDQDFDLSQLHRGL 767
QY 723 EAREPVVLRNDVAPTIPTPMYPRRANPDEIGNFTIENKAANTDPTAPPYDTLLVFDY 782
DB 768 DARPDVT-RNDVAP-VLAAPQYPRRANPDEIGNFTIDENLNAADNDPTAPPYDSLLVFDY 825
QY 783 EGSGSAAASLSSTTSASDQDQDYVLNWSGPFKPLADMYGGED 828
DB 826 EGSGSAAASLSLSNNSDLDQDYSALNDWNGPRFTKLADMYGGED 871

RESULT 12
IJHUCN

cadherin 2 precursor - human
N;Alternate names: N-cadherin; neuronal cadherin
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C;Accession: A38870; S11487; JQ0751; S13799
R;Reid, R.A.
submitted to the EMBL Data Library, November 1990
A;Reference number: A38870
A;Accession: A38870
A;Molecule type: mRNA
A;Residues: 1-906 <RE1>
A;Cross-references: UNIPROT:PI9022; UNIPARC:UPI000016ADBE; EMBL:X54315; NID:G34998; PIDN
R;Reid, R.A.; Hamperly, J.J.
Nucleic Acids Res. 18, 5896, 1990
A;Title: Human N-cadherin: nucleotide and deduced amino acid sequence.
A;Reference number: S11487; MUID:91016946; PMID:2216790
A;Accession: S11487
A;Molecule type: mRNA
A;Residues: 1-340,'N',342-698,'R',700-704,'F',706-906 <RE2>
A;Cross-references: UNIPARC:UPI000017434F; EMBL:X54315
A;Note: this sequence has been revised in reference A38870
R;Walsh, F.S.; Barton, C.H.; Putt, W.; Moore, S.E.; Kelsell, D.; Spurr, N.; Goodfellow, I.
J. Neurochem. 55, 805-812, 1990
A;Title: N-cadherin gene maps to human chromosome 18 and is not linked to the E-cadherin
A;Reference number: JQ0751; MUID:90347462; PMID:2384753
A;Accession: JQ0751
A;Molecule type: mRNA
A;Residues: 160-194,'IR',197-211,'L',213-227,'Q',229,'N',231-235,'G',237-248,'T',250-356;
A;Cross-references: UNIPARC:UPI000016ADCO; GB:M34064
C;Comment: Cadherins mediate calcium-dependent intercellular adhesion and are thought to
C;Genetics:
A;Gene: GDB:CDH2; NCAD
A;Cross-references: GDB:128185; OMIM:114020
A;Map position: 18q12.1-18q12.1
C;Superfamily: cadherin; cadherin repeat homology
C;Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane prot
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-159/Domain: propeptide #status predicted <PRO>
F;160-906/Product: N-cadherin #status predicted <MAT>
F;160-714/Domain: extracellular #status predicted <EXT>
F;162-267/Domain: cadherin repeat homology <CR1>
F;217-242/Region: cadherin repeat #status predicted
F;270-382/Domain: cadherin repeat homology <CR2>
F;385-497/Domain: cadherin repeat homology <CR3>
F;500-605/Domain: cadherin repeat homology <CR4>
F;606-712/Domain: cadherin repeat homology <CR5>
F;715-746/Domain: transmembrane #status predicted <TM>
F;747-906/Domain: intracellular #status predicted <INT>
F;865-878/Region: serine-rich
F;190,273,325,402,572,622,651,692/Binding site: carbohydrate (asn) (covalent) #status pre

Query Match 39.1%; Score 1707; DB 1; Length 906;
Best Local Similarity 42.7%; Pred. No. 9.6e-101;
Matches 390; Conservative 129; Mismatches 292; Indels 102; Gaps 27;

QY 6 GPLASLLLLQVCQAASEP-----CRAVFEAEVLEAGAEQPGQALGKV-FWGC 59
DB 6 GALTLLPLLALLQASVEASGEIALCKTGPPE-DVYSAVLSKDVHEGQPLNVPKFSNCN 64
QY 60 GO-----EPALSTDNDD--FTVRN-----GETVQSR----- 84
DB 65 GKRKVQYESSEPADFKVDEGDMVAVRSPFLSEHAKFLIYAQDKETQKQVAVKLSLK 124
QY 85 -----RSLKERNPLK--IPE-----SKRI--LRHRKRDWVAVIPGVPENGKGPFPORLNOLK 132
DB 125 PTLTEESVKESAEEVEIIVFPQFSKSHGLQKQKRDWVIPPINLPENSGPFPQELVLR 184
QY 133 SNKDRDTKIFYSTGPGADSPPEGVFAVEKETGMLLNKPLDREETAKYELFGHVSNG 192
DB 185 SDRDKNLSURYTVTGTGADQPTGTGIINPIISQLSVTKPLDREQIARPHLRAHAYDING 244

Qy 193 ASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLGVLPSTVMOVATDDEDAIITYNGV 252
Db 245 NOVENPDIIVINVIDMNDNRPEFLHQVWNGTVPESKEFGYVMTVTAIDADDP-NALNGM 303
Qy 253 VAKSIHQEPKDPHLMFTIHRSTGTISVISGGLDREKVPYETLTIIQATDMOG---DGST 309
Db 304 LRYRIVSQAPSTPSNMFNTINNETGDIITVAAGLDREKVQOYTLIIQATDMEGIPTVGLS 363
Qy 310 TTAVALVEILDANDANPMFDPQKYEAHVPENAVGHEVQRLTVTDLDPNSPAMRATYLM 369
Db 364 NTATAVITVTDVNDNPEFTAMTVEYGPENRVDIIIVANLTVTDKQPHTPAMNAVYRIS 423
Qy 370 GDDGDGHTITTHPESNOGILTRKGLDFFAKNQHTLYVEVNEAPVFLXL---PTSTAT 426
Db 424 GGDPTGFALQTDNSNDGLVTVVKPIDFTNRMFVLTVAAENQVFLAKGIQHPQSTAT 483
Qy 427 IVRVHVEDVNEAPVFPSPKVVVEQEGIPTGPEVCVYTAEDPK-ENQKISYRILRDPAGW 485
Db 484 VSVTVIDVNEPYPAPNPKIIRQEBGLHAGTMLTFTTAQDPDRYMQQNIYTKLSDFANW 543
Qy 486 LAMPDPSGOVAVTGLDREDEQFVRNNIYEVWVLAMONGSPPTGTGTLLTLTLIDVNDHG 545
Db 544 LKIDVNGQITTIATVLDRESN-VKNNIYNATFLASDNGIPPMSTGTTLQIYLLDINDNA 602
Qy 546 PVPEPRQITICNOSPVRHVLNIT--DKDLSPTSPFOAQLTDDSDIY---WTAEVNEEGD 600
Db 603 PQVLPQBAETC-ETPDPSINITALDYDIDNAGPFADPLPSVTTIKRWNTI--TRLNGD 660
Qy 601 TVVLISLK-KFLQDQTYDVHLSLSDHGN--KEQLTVIRATVCDCHG-----VETCPGPWK 652
Db 661 PAQLNLKIKFLKLEAGIYEVPIIITDSGNPPKSNISILRVKVCQCSNGDCTDVDRIVGAGL 720
Qy 653 GGFILPVLGAVLALLFLLVLLLV-----RKCRKIKEPLLLPEDDTRDNVYFG 702
Db 721 G-----TGAIATALLCIIILLVLMFVVMKRDKERQAKQLLIDPEDDVRDNLKYD 774
Qy 703 BEGGGEEDQDYDTQLHRGLEARPEV-----VLKNDVAPTIIPTPMYRPPRA--NPDEIG 755
Db 775 BEGGGEEDQDYDLSQLQDPDVEADAIPKVGIRMDERP-IHAEPQYFVRSAAHPGDIG 833
Qy 756 NFIIENKAANTDPTAPPYDYLVDYEGSGDAASSLTSSASDQDQDYLYLNGSR 815
Db 834 DFINEGLKAANDPTAPPYDYLVDYEGSGTAGSLSSLSGSGGEGQDYDYLYLNDGPR 893
Qy 816 FKKLADMYGGGED 828
Db 894 FKKLADMYGGGDD 906

RESULT 13
tUCHCN
N-cadherin precursor, neuronal - chicken
C:Species: Gallus gallus (chicken)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C:Accession: A29964
R:Hatta, K.; Nose, A.; Nagafuchi, A.; Takeichi, M.
J. Cell Biol. 106, 873-881, 1988
A:Title: Cloning and expression of cDNA encoding a neural calcium-dependent cell adhesion molecule
A:Reference number: A29964; PMID:88153917; PMID:2831236
A:Accession: A29964
A:Molecule type: mRNA
A:Residues: 1-312 <HAT>
A:Cross-references: UNIPROT:P10288; UNIPARC:UPI00001269B8; GB:X07277; NID:g36349; PIDN:C
C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-164/Domain: propeptide #status predicted <PRO>
F:165-912/Product: N-cadherin #status predicted <MAT>
F:167-272/Domain: extracellular #status predicted <EXT>
F:242-247/Region: cadherin binding #status predicted

F:275-387/Domain: cadherin repeat homology <CR2>
F:390-502/Domain: cadherin repeat homology <CR3>
F:505-611/Domain: cadherin repeat homology <CR4>
F:612-720/Domain: cadherin repeat homology <CR5>
F:721-752/Domain: transmembrane #status predicted <TMM>
F:753-912/Domain: intracellular #status predicted <INT>
F:869-884/Region: serine-rich
F:278,330,407,578,628,657/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 39.0%; Score 1704.5; DB 1; Length 912;
Best Local Similarity 42.8%; Pred. No. 1.4e-100;
Matches 390; Conservative 132; Mismatches 284; Indels 109; Gaps 27;

Qy 7 PLASLILL--LQVCWLOCAASEP--CRAVFRE----AEVTLKAGAEQEPGQALGV-FMGC 58
Db 14 PLALMLLALQAPAKATCEDMLCKMGFPEDVHSAVVSRSVHG-----GQPLLNVRFQSC 68
Qy 59 PGQEPALF-STDNDDFTV-----RNGETVQERR----- 85
Db 69 DENRKIVFGSSEPEDFRVGDGVVYAERSFQLSAEPTEFVVSARDKETQBEWQMKVLTLP 128
Qy 86 -----SLKERNPLK--IPPSKRI-----LRRHKRDVVVAPISVPENGGKPPQRLNQL 131
Db 129 EPAFTGASEKQDKKIEDIIPWQOYKXSSHLKQKQKDWIIPPINLPENSRGPPQELVRI 188
Qy 132 KSNKDRDTKIFYGITGPGADSPPEGVFAVEKETGWLKLNKPLDREETAKYELFGHVAESEN 191
Db 189 RSRDRLSLRYSVTGEGADQPTGFIINPISQSLSVTKPLDREQLASFLHRAHADVN 248
Qy 192 GASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLGVLPSTVMOVATDDEDAIITYNG 251
Db 249 GNQVENPIDIVINVIDMNDNRPEFLHQVWNGTVPESKPGTYVMTVTAIDADDP-NAQNG 307
Qy 252 VVAYSHTSQBPDPHLMFTIHRSTGTISVISGGLDREKVPYETLTIIQATDMOGD---GS 308
Db 308 MLRYRILSQAPSPSPNMFNTINNETGDIITVAAGLDREKVQOYTLIIQATDMENPTYGL 367
Qy 309 TTTAVAVVEILDANDANPMFDPQKYEAHVPENAVGHEVQRLTVTDLDPNSPAMRATYLI 368
Db 368 SNTATAVITVTDVNDNPEFTAMTVEYGPENRVDIIIVANLTVTDKQPHTPAMNAVYQM 427
Qy 369 MGGDGDGHTITTHPESNOGILTRKGLDFFAKNQHTLYVEVNEAPVFLXL---PTSTA 425
Db 428 TGGDPTGQFTILTDPNSNDGLVTVVKPIDFTNRMFVLTVAAENQVFLAKGIQHPQSTA 487
Qy 426 TIVVHVEDVNEAPVFPSPKVVVEQEGIPTGPEVCVYTAEDPKENOKIS--YRILRDP 483
Db 488 TVSITVIDVNESPYFVNPVKLVQEBGLLAGSMLTFTTARDPDRYMQQTSURYSKLSDPA 547
Qy 484 GWLAMPDPSGOVAVTGLDREDEQFVRNNIYEVWVLAMONGSPPTGTGTLLTLTLIDVND 543
Db 548 NLWKIDVPNGQITTVAVLDRE-STYVQNNYNNATFLASDNGIPPMSTGTTLQIYLLIND 606
Qy 544 HGPVPEPRQITICNOSPVRHVLNIT--DKDLSPTSPFOAQLTDDSDIY---WTAEVNEE 598
Db 607 NAPQVNEKATTC-ETLQPNAINITAVDPIDPNAGPFAPELDPSPSIRKNNMTI-VRIS 664
Qy 599 GDTVVLSLK-KFLQDQTYDVHLSLSDHGN--KEQLTVIRATVCDCH-----GHVETCEGP 650
Db 665 GDHAQLSLRIFLEAGIYDYPVITDSGNPHASSTSVLKVKVCQCDINGDCTDVRIVGA 724
Qy 651 WKGGFILPVLGAVLALLFLLVLLLV-----RKCRKIKEPLLLPEDDTRDNVVFY 700
Db 725 GLG-----TGAIATALLCIIILLVLMFVVMKRDKERQAKQLLIDPEDDVRDNLK 778
Qy 701 YGEGGGEEDQDYDTQLHRGLEARPEV-----VLKNDVAPTIIPTPMYRPPRA--NPDE 753
Db 779 YDESGGGEEDQDYDLSQLQDPDVEADAIPKVGIRMDERP-IHAEPQYFVRSAAHPGD 837
Qy 754 IGFIENLKAANTDPTAPPYDYLVDYEGSGDAASSLTSSASDQDQDYLYLNGW 813
Db 838 IGDFINEGLKAANDPTAPPYDYLVDYEGSGTAGSLSSLSGSGGEGQDYDYLYLNDWG 897


```
QY 814 SRFKKLADMYGGGED 828
Db 898 PRFKKLADMYGGGDD 912

RESULT 14
IJM5CN
N-cadherin precursor, neuronal - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
R;Accession: A32759; A46163
R;Miyatani, S.; Shimamura, K.; Hatta, M.; Nagafuchi, A.; Nose, A.; Matsunaga, M.; Hatta,
Science 245, 631-635, 1989
A;Title: Neural cadherin: role in selective cell-cell adhesion.
A;Reference number: A32759; MUID:89346748; PMID:2762814
A;Accession: A32759
A;Molecule type: mRNA
A;Residues: 1-906 <MI2>
A;Cross-references: UNIPROT:P15116; UNIPARC:UPI000002924B; GB:M31131; NID:g192327; PIDN:
R;Miyatani, S.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Takeichi, M.
Proc. Natl. Acad. Sci. U.S.A. 89, 8443-8447, 1992
A;Title: Genomic structure and chromosomal mapping of the mouse N-cadherin gene.
A;Reference number: A46163; MUID:92409532; PMID:1528849
A;Accession: A46163
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 839-906 <MI2>
A;Cross-references: UNIPARC:UPI0000170C39; GB:S45011; NID:g256010; PIDN:AAB23356.1; PID:
A;Note: sequence extracted from NCBI backbone (NCBIN:113759, NCBI:P.113760)
C;Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
C;Superfamily: cadherin; cadherin repeat homology
F;1-27/Domain: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-159/Domain: propeptide #status predicted <PRO>
F;160-906/Product: N-cadherin #status predicted <WAT>
F;160-906/Domain: extracellular #status predicted <EXT>
F;162-267/Domain: cadherin repeat homology <CR1>
F;237-242/Region: cadherin binding #status predicted
F;270-382/Domain: cadherin repeat homology <CR2>
F;385-497/Domain: cadherin repeat homology <CR3>
F;500-605/Domain: cadherin repeat homology <CR4>
F;606-714/Domain: cadherin repeat homology <CR5>
F;715-746/Domain: transmembrane #status predicted <TMW>
F;747-906/Domain: intracellular #status predicted <INT>
F;865-978/Region: serine-rich
F;190,273,325,402,572,651,692/Binding site: carbohydrate (Asn) (covalent) #status predic
```

```
Db 306 YRILSQAPSTPSPNMTINNETGDIITVAAGLDREKQQVTLIIQATMEGNETYGLSNT 365
QY 312 AVAVVEILDANDNAPMFDQPKTEAHVPENAVGHEVQRLTVTDLDAFNSPAWRATYILMGG 371
Db 366 ATAVITVDNDNPPEFTAMTFYGEVPENRVDIVANLTVTDKQPHPTFAMNAAYRISGG 425
QY 372 DGDHFTIITHPESNOGILTTKGLDFEAKOHTLVVEVTNEAPFVLKL---PTSTATIV 428
Db 426 DPTGRFAILTDPNSNDGLTVVVKPIDFETNRMFVLTVAAENQVPLAKGIQHPPQSTATVS 485
QY 429 VHVEDVNEAPVFPVPSKVVEVEGIGTGPVCVYTAEDPK-ENOKISVRIILDRPAGWLA 487
Db 486 VTVIDVNEPYPAPNDKIRQEBGLHAGTWLTLTAQDDPRYMQQNRITKLSDPANWLK 545
QY 488 MDPDSQVTAAGTLDREDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLTLDVNDHGPV 547
Db 546 IDPVNGQITIAVLDRS-SPYVQNNIYNATFLASDNGIIPMSCTGTLOIYLLDINDNAPQ 604
QY 548 PEPRQITICNQSPVRHVLNIT--DKDLSPTSPFQAQLTDDSDIY---WTAEVNE-EGDT 601
Db 605 VLPQEAETC-ETPEPNSINIAALVDIDPNAGPFAPDPLSPVTIKRNT--INRLNGDF 661
QY 602 VVLSLK-KELQDQTYDVHLSLSDHGN--KEQLTVIRATVCDCHG-----VETCPGPWK 653
Db 662 AQLNLKIKLEAGIYEVPIIITDSGNPKNSILRVKVCQDSNGDCTDVIDRIVGAGLG 721
QY 654 GTPLVGLGAVLALLFLLLVLLLV-----RKRKIKEPLLLPEDDTRONVIFYGE 703
Db 722 -----TGAIILLCIIILLILVFMVVMKRDKERQAKOLLIDPEDDVRDNLKYDE 775
QY 704 EGGEBDQYDITQLHARGLEAREV-----VLKNDVAPITPTPMYRPPA--NPDEIGN 756
Db 776 EGGEBDQYDLSQLQPDVTVEPDAIKPVGIRRLDERP-IHAEPQYFVRSAAHPHGDIGD 834
QY 757 FIENLKAANTOPTAPYDTLAVEDYEGSGSDAASLSLTSASDQDQYDYLNENGSRF 816
Db 835 FINEGLKAADNPTAPPYDSLLVFDYEGSGTAGSLSSLNSSSGGQDQDYLDNDGPRF 894
QY 817 KKLADMYGGGED 828
Db 895 KKLADMYGGGDD 906

RESULT 15
IJCHCR
R-cadherin precursor - chicken
C;Species: Gallus gallus (chicken)
C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
C;Accession: JH0424
R;Inuzuka, H.; Miyatani, S.; Takeichi, M.
Neuron 7, 69-79, 1991
A;Title: R-cadherin: a novel Ca2+-dependent cell-cell adhesion molecule expressed in the
A;Reference number: JH0424; MUID:91299341; PMID:1712604
A;Accession: JH0424
A;Molecule type: mRNA
A;Residues: 1-913 <INU>
A;Cross-references: UNIPROT:P24503; UNIPARC:UPI0000126D94; GB:D14459; GB:D00849; NID:g22;
A;Experimental source: retina
C;Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
C;Superfamily: cadherin; cadherin repeat homology
C;Keywords: calcium binding; cell adhesion; duplication; glycoprotein; retinal; transmembr
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-166/Domain: propeptide #status predicted <PRO>
F;167-913/Product: R-cadherin #status predicted <WAT>
F;167-913/Domain: extracellular #status predicted <EXT>
F;169-274/Domain: cadherin repeat homology <CR1>
F;244-249/Region: cadherin binding #status predicted
F;277-389/Domain: cadherin repeat homology <CR2>
F;392-504/Domain: cadherin repeat homology <CR3>
F;507-612/Domain: cadherin repeat homology <CR4>
F;613-721/Domain: cadherin repeat homology <CR5>
F;722-753/Domain: transmembrane #status predicted <TMW>
F;754-913/Domain: intracellular #status predicted <INT>
```

Job time : 48 secs

F;870-885/Region: serine-rich
F;280,409,554,629,658,699/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match		38.8%;	Score 1694;	DB 1;	Length 913;
Best Local Similarity		41.5%;	Pred. No. 6.5e-100;		
Matches 383;		Conservative 135;	Mismatches 286;	Indels 118;	Gaps 25;
Qy	11	LLLLQVCWLQCAA-----SEPCRAVFRAEAVLEAGGAEQEPGQALGKV-FMGCPGOE	62		
Db	7	LLLLVLLVWGSAALNGDLTVRPTCKPGFSBEDYTAFSQNIIE-GQKLLKVKFNKNCAGNK	65		
Qy	63	PALFSTDNDPFTVR-----	76		
Db	66	GVRVETNSLDFKVRADGTWVAHVQVMASKQLILMVTAWDPQTLGRWEAIVRFLVGEKLQ	125		
Qy	77	-NGETVQERRS-----LKERNPKLFP-----SKRILRRHKRDWVADISVPENCKGPFPO	126		
Db	126	HNGHKPKGRKSGPVDLAQQQSDTLTPWRQHSAKGLRQKRDWVIPPINVPENSRGPFPO	185		
Qy	127	RLNOLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDRREEIAKYELFGH	186		
Db	186	QLVRIKDKKEIHRISITGVGADQDPMEVFSIDPVGSRMYVTRPMDREERASYHLRAH	245		
Qy	187	AVSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGSLGVLPGTSVMQVTTATDEDDAI	246		
Db	246	AVDMNGNKVENPIDLYIYVVIDMNDNRPEFINQVYNGSVDEGSKPGTYVMTVTANDADST	305		
Qy	247	YTYNGVAVYSIHSQEPKDPHDLMETIHRSTGTSIVISSGLDREKVPETLTIOATDMDGD	306		
Db	306	-TANGMVRYRIVTQTPQSPNQNMFTINSETGDIYVTAAGLDREKVQOQVMIVQATDMEGN	364		
Qy	307	--GSTTAVAVRILDANDNAPMFDQKYEAHVPENAVGEVQRLTVTDLDAENSPAWR	363		
Db	365	LNYSGLSWTATAIIVTVDNDNPPEFTTSTYSGEVPENRVEVVVANLIVMDRDQPHSPWN	424		
Qy	364	ATYLMGGDDGDHFTITTHPESNQGIITFRKGLDFEAKNQHTLYVEVTNEAPFVLKLP--	421		
Db	425	AIYRIISGDPGSGHTTIRTPVTNEGTVVKAVDYEMNRAFMLTVMVSNQAPLASGIQMS	484		
Qy	422	-TSTATIVVHVEDVNEAPVFPSPKVEVEGIGITGEPVCVYTAEDPDK-ENKISYRIL	479		
Db	485	FQSTAGVTISVTDVNEAPYFPTNHKLIRLEEGVPTGTLTTFSAVDPRFMQQAARYSKL	544		
Qy	480	RDPAWGLAMDPSQVTAAGTLDREDOFVRNNIYEVWVLAMDNGSPPTGTGTLLTLTI	539		
Db	545	SDPANWLNINATNGQITTAAVLDRESD-YIKNNVYEATFLAADNGIPPASGTGTLQIYLI	603		
Qy	540	DVNDHGPVPEPQITICNQSPVRHVNIT--DKDLSPHTSPPQAQLTDDSDIY---WTAE	594		
Db	604	DINDNAPELLPEKAQIC-EKENLVNINITAADADIDPNVGPFPPELPSPSAVRKNWTI-	661		
Qy	595	VNEEGDTVVLISLK-KFLKQDITYDHLSDHGNKE--QLTVIRATVCDG--HGHVETCPG	649		
Db	662	TRLNGDYAQLSLRMYLEAGYVDYPIIIVTDSGNPPLYNTSIIKVKVCPCDENGDCITIGA	721		
Qy	650	PWKGGFILPVLGAVLALLFLLLVLLV-----RKKRKIKIPEPLLPEDDDTRDNVF	699		
Db	722	VAAAGL---GTGAIAIILICIIILITWVLLFVWWMKREKERHTKQLIIDPEDDVRNII	778		
Qy	700	YYGEGGEEEDQDYDITQLHRGLEARPEV-----VLRNDVAPTIITPTMYRPRPA	749		
Db	779	KYDEEGGEEEDQDYDLSLQ-----QPETMDHVLNKAPGVRVRVDERP-IGAEPQYPIRPV	832		
Qy	750	--NDEIGNFIENLKAANTDPTAPPYDTLLVFDYEGSGSDAASLSLTSSASDQDQDYD	807		
Db	833	IPHEGDIIGDFINEGLRAADNDPTAPPYDLSLVFDYEGSGTAGSVSLNSSSG-DQDYD	891		
Qy	808	YLNWGSRRFKKLADMYGGGEDD	829		
Db	892	YLDNWGPRFKKLADMYGGGEED	913		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2005, 09:01:57 ; Search time 189 Seconds
(without alignments)
1927.222 Million cell updates/sec

Title: US-10-696-639-39

Perfect score: 4369

Sequence: 1 MGLPRGLASLLLLQVCMQLQ.....NEWGSRFKLADMYGGGSD 829

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4369	100.0	829	5	ABG61897	Abg61897 Prostate
2	4369	100.0	829	5	ABb81476	Abb81476 Human P-c
3	4369	100.0	829	5	ABj05598	Abj05598 Breast ca
4	4369	100.0	829	5	AAu97492	AAu97492 Human P-c
5	4369	100.0	829	5	AAm50864	AAm50864 Cadherin
6	4369	100.0	829	6	ABP58357	ABP58357 Human P-c
7	4369	100.0	829	7	ADc15497	ADc15497 Human bas
8	4369	100.0	829	7	ADd14190	ADd14190 Human src
9	4369	100.0	829	7	ADn39482	ADn39482 Cancer/an
10	4369	100.0	829	7	ADn39578	ADn39578 Cancer/an
11	4369	100.0	829	7	ADn39545	ADn39545 Cancer/an
12	4369	100.0	829	7	ADn39465	ADn39465 Cancer/an
13	4369	100.0	829	8	ADL06565	ADL06565 Human tum
14	4369	100.0	829	8	ADQ20015	ADQ20015 Human sof
15	4369	100.0	829	8	ADP26905	ADP26905 Human P-c
16	4369	100.0	829	9	ADV73198	ADV73198 Human col
17	4369	100.0	829	9	AEb87752	AEb87752 Human P-c
18	4361	99.8	829	5	ABP54683	ABP54683 Metastati
19	4361	99.8	829	6	ABR58670	ABR58670 Human can
20	4361	99.8	829	6	ABU56434	ABU56434 Lung can
21	4361	99.8	829	7	ADN39020	ADN39020 Cancer/an
22	4361	99.8	829	7	ADN39928	ADN39928 Cancer/an
23	4361	99.8	829	8	ADL70229	ADL70229 Colon can
24	4361	99.8	829	8	ADN59621	ADN59621 Colon neo

25	4361	99.8	829	8	ADN03948	Adn03948 Antipeori
26	4361	99.8	829	9	ADO28657	Ado28657 Human CAD
27	4361	99.8	829	9	ADx83229	Adx83229 Human TEG
28	4351	99.6	829	6	ABU56670	Abu56670 Lung can
29	3643.5	83.4	700	8	ADP29426	Adp29426 Human sec
30	3643.5	83.4	700	8	ADP29333	Adp29333 Human sec
31	2420	55.4	882	2	AAI09375	Aay09375 Wild-type
32	2420	55.4	882	3	AAH35730	Aab35730 Human E-c
33	2420	55.4	882	4	AAH73490	Aab73490 Human E-c
34	2420	55.4	882	5	ABG96293	Abg96293 Human ova
35	2420	55.4	882	5	AAU78051	Aau78051 Human E-c
36	2420	55.4	882	5	ADL15660	Adl15660 Human E-c
37	2420	55.4	882	8	ADN03700	Adn03700 Antipeori
38	2420	55.4	882	8	ADP26906	Adp26906 Human E-c
39	2420	55.4	882	8	ABM81765	Abm81765 Tumour-as
40	2420	55.4	882	8	ADr67260	Adr67260 Human bla
41	2420	55.4	882	9	AEb87740	Aeb87740 Human E-c
42	2420	55.4	896	8	ADP66291	Adp66291 Human pro
43	2420	55.4	896	8	ADR66852	Adr66852 Human pro
44	2420	55.4	923	8	ADR66668	Adr66668 Human pro
45	2420	55.4	923	8	ADR66326	Adr666326 Human pro

ALIGNMENTS

RESULT 1
ABG61897
ID ABG61897 standard; protein; 829 AA.
XX
AC ABG61897;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #98.
DE
KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS Mammalia.
XX
PN WO200230268-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US032045.
XX
PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733288.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 06-APR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
XX
(EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
XX WPI; 2002-471335/50.
XX N-PSDB; ABK92214.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
prostate tissue.
XX
PS Claim 27; Page 384-385; 436pp; English.
XX
XX The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises

102 a

CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
XX

SQ Sequence 829 AA;

Query Match 100.0%; Score 4369; DB 5; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGLASLLLLQVLCWQCAASEPCRAVFAEVTLKAGGAEQEPQALGKVFMCPCG 60
Db 1 MGLPRGLASLLLLQVLCWQCAASEPCRAVFAEVTLKAGGAEQEPQALGKVFMCPCG 60

Qy 61 QEPALFSTDNDFTVRNGETVOERRSLKERNPLKIFPSKILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDFTVRNGETVOERRSLKERNPLKIFPSKILRRHKRDWVAPISVPENG 120

Qy 121 KGPPFQRLNQLKSKNRDRTKIFYGITGPGADSPPEGVFAVEKETGWLNNKPLDREIAK 180
Db 121 KGPPFQRLNQLKSKNRDRTKIFYGITGPGADSPPEGVFAVEKETGWLNNKPLDREIAK 180

Qy 181 YELFGHAVSENGASVEDPMNLSIIVTDQNDHKPKFTQDTFRGSVLEGVLPCTSYMQVTAT 240
Db 181 YELFGHAVSENGASVEDPMNLSIIVTDQNDHKPKFTQDTFRGSVLEGVLPCTSYMQVTAT 240

Qy 241 DEDDAITYNGVWVAYSISQBPDPKPHDLMFTIHRSTGTISVISGLDREKVPETLTIQA 300
Db 241 DEDDAITYNGVWVAYSISQBPDPKPHDLMFTIHRSTGTISVISGLDREKVPETLTIQA 300

Qy 301 TDMGDGSGTTTAVAVVEILDANDNAMPDPQKYEAHPVENAVGHEVQRLTVDLADPNSP 360
Db 301 TDMGDGSGTTTAVAVVEILDANDNAMPDPQKYEAHPVENAVGHEVQRLTVDLADPNSP 360

Qy 361 AWRTATYLINGDGDGHFTITTHPSNQILITRGLDFEAKNQHTLYVEVNEAPFVLKL 420
Db 361 AWRTATYLINGDGDGHFTITTHPSNQILITRGLDFEAKNQHTLYVEVNEAPFVLKL 420

Qy 421 PTSTATIIVHVEDVNEAPVFPVPPSKVVEVQGIPTGEPVCVYTAEDDPKENOKISYRILR 480
Db 421 PTSTATIIVHVEDVNEAPVFPVPPSKVVEVQGIPTGEPVCVYTAEDDPKENOKISYRILR 480

Qy 481 DPAGWLAMPDPSGQVTAAGTLTLDREDEQFVRNNIYEVWVLAMDNQSPPTGTGTLTLLTID 540
Db 481 DPAGWLAMPDPSGQVTAAGTLTLDREDEQFVRNNIYEVWVLAMDNQSPPTGTGTLTLLTID 540

Qy 541 VNDHGVPPEPQITICNSPVRHVLNITDKLSPTSPFQALTDSDSIYTAEVNEEGD 600
Db 541 VNDHGVPPEPQITICNSPVRHVLNITDKLSPTSPFQALTDSDSIYTAEVNEEGD 600

Qy 601 TVVLSLKKFLKQDVTVDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFLLPVL 660
Db 601 TVVLSLKKFLKQDVTVDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFLLPVL 660

Qy 661 GAVLALLFLLVLLVLRKGRKIKEPPLLPEDDTRDNVFFYGEESGGEDQDITQLHR 720
Db 661 GAVLALLFLLVLLVLRKGRKIKEPPLLPEDDTRDNVFFYGEESGGEDQDITQLHR 720

Qy 721 GLEARPEVLRNDVAPTIIPTPMTRPRPANPDEIGNFIENLKAANTDPTAPPDITLLVF 780
Db 721 GLEARPEVLRNDVAPTIIPTPMTRPRPANPDEIGNFIENLKAANTDPTAPPDITLLVF 780

Qy 781 DYEGSGSDAASLSLTSASDQDDYDLNENWGSFRFKLADMYGGEDD 829

Db 781 DYEGSGSDAASLSLTSASDQDDYDLNENWGSFRFKLADMYGGEDD 829

RESULT 2
ABBS1476

ID ABB81476 standard; protein; 829 AA.

XX ABB81476;

XX 30-AUG-2002 (first entry)

XX Human P-cadherin protein SEQ ID NO:50.

XX Cadherin; alpha-catenin; cancer; beta-catenin binding domain; melanoma;
XX O-catenin; colon cancer.

XX Homo sapiens.

XX US2002045591-A1.

XX 18-APR-2002.

XX 17-JUL-2001; 2001US-00905983.

XX 26-MAY-1998; 98IL-00124650.

XX 26-MAY-1999; 99US-00318633.

XX (GEIG/) GEIGER B.

XX (BENZ/) BEN-ZE'EV A.

XX (SADO/) SADOT E.

XX Geiger B, Ben-Ze'ev A, Sadot E;

XX WPI; 2002-499105/53.

XX N-PSDB; ABN89393.

XX New construct encoding soluble cytoplasmic portion of cadherin including
XX beta catenin binding domain useful in treating cancer associated with
XX high beta-catenin activity e.g. colon cancer and melanoma.

XX Example 3; Page 53-55; 102pp; English.

XX The present invention describes a pharmaceutical composition for treating
XX cancer associated with abnormally high beta-catenin activity. The
XX pharmaceutical composition comprises a gene therapy vehicle harbouring a
XX polynucleotide that contains: (i) a nucleotide sequence encoding a
XX soluble cytoplasmic portion of a cadherin which lacks a transmembrane
XX portion and an extracellular portion of the cadherin, and includes a beta
XX -catenin binding domain; and (b) an upstream promoter for directing
XX expression of the soluble cytoplasmic portion of the cadherin in a
XX mammalian cell. Also described is a pharmaceutical composition for
XX treating cancer associated with abnormally high activity levels of beta-
XX catenin comprising a gene therapy vehicle harbouring a polynucleotide
XX that contains: (a) a nucleotide sequence encoding an O-catenin; and (b)
XX an upstream promoter for directing expression of the O-catenin in a
XX mammalian cell. The pharmaceutical compositions have cytostatic activity
XX and can be used in the suppression of beta-catenin-mediated
XX transactivation. They can be used for treating cancers associated with
XX abnormally high activity levels of beta-catenin such as colon cancers and
XX melanomas, by reducing these high activity levels of beta-catenin in
XX mammalian cells. The present sequence represents human P-cadherin which
XX is used in the exemplification of the present invention

SQ Sequence 829 AA;

Query Match 100.0%; Score 4369; DB 5; Length 829;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGLASLLLLQVLCWQCAASEPCRAVFAEVTLKAGGAEQEPQALGKVFMCPCG 60

Db 1 MGLPRGLASLLLLQVLCWQCAASEPCRAVFAEVTLKAGGAEQEPQALGKVFMCPCG 60

QY	61	QEPALFSTNDNDDFTVRNGETVQERRSLKERNPLKIFPPSKRILRRHKRDVWVAPISVPENG	120
Db	61	QEPALFSTNDNDDFTVRNGETVQERRSLKERNPLKIFPPSKRILRRHKRDVWVAPISVPENG	120
QY	121	KGPPFORLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLINLKLPLDREI	180
Db	121	KGPPFORLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLINLKLPLDREI	180
QY	181	YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSLVGLPGTSGVMQV	240
Db	181	YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSLVGLPGTSGVMQV	240
QY	241	DEDDAIYTYNGVAVYSIHQEPKDPHDLMTFTHRSTGTISVSSGLDREKVPYETL	300
Db	241	DEDDAIYTYNGVAVYSIHQEPKDPHDLMTFTHRSTGTISVSSGLDREKVPYETL	300
QY	301	TMMDGSGSTTTAVAVVEILDANDNAPMFPDQKYEAHVPENAVGVHVEVQRLTV	360
Db	301	TMMDGSGSTTTAVAVVEILDANDNAPMFPDQKYEAHVPENAVGVHVEVQRLTV	360
QY	361	AWRATYLMGGDDGDHFTTTHPESNOGILTRKGLDFAKQHTLYVEVTNEAPFV	420
Db	361	AWRATYLMGGDDGDHFTTTHPESNOGILTRKGLDFAKQHTLYVEVTNEAPFV	420
QY	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEQGIPTGEPVCVYTAEDPDKENQKIS	480
Db	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEQGIPTGEPVCVYTAEDPDKENQKIS	480
QY	481	DPAGWLAMPDQSGQVAVGLDREDFQVRNNIYEVWVLMONGSPPTTGTGLTL	540
Db	481	DPAGWLAMPDQSGQVAVGLDREDFQVRNNIYEVWVLMONGSPPTTGTGLTL	540
QY	541	VNDHGVPPEPRQITICNQSPVRHLNITDKDLSPHTSPPQAOLTDSDIYWTAEVNE	600
Db	541	VNDHGVPPEPRQITICNQSPVRHLNITDKDLSPHTSPPQAOLTDSDIYWTAEVNE	600
QY	601	TVVLSLKKFLKQDYDVHLSDHGNKEQLTVIRATVCDCHGVETCPGPKGFFLP	660
Db	601	TVVLSLKKFLKQDYDVHLSDHGNKEQLTVIRATVCDCHGVETCPGPKGFFLP	660
QY	661	GAVLALLFLLVLLVLRKRIKEPLLPEDDTRDNVIFYEGEGGEDDYDITQLHR	720
Db	661	GAVLALLFLLVLLVLRKRIKEPLLPEDDTRDNVIFYEGEGGEDDYDITQLHR	720
QY	721	GLEARPEVLRNDVAPTIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYD	780
Db	721	GLEARPEVLRNDVAPTIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYD	780
QY	781	DYEGSGSDAASLSSTSSASDQDDQDYDLNMGSRFKKLADMYGGGEDD	829
Db	781	DYEGSGSDAASLSSTSSASDQDDQDYDLNMGSRFKKLADMYGGGEDD	829

RESULT 3

ABJ05598
ID ABJ05598 standard; protein; 829 AA.

XX
AC ABJ05598;

XX
DT 14-NOV-2002 (first entry)

XX
DE Breast cancer-associated protein 63.

XX
KW Breast cancer; breast cancer-associated gene sequence; drug development;
pharmacogenetics; biosensor development.

XX
OS Unidentified.

XX
PN WO200259377-A2.

XX
PD 01-AUG-2002.

XX

24-JAN-2002; 2002WO-US002242.

24-JAN-2001; 2001US-0263965P.
02-FEB-2001; 2001US-0265928P.

09-APR-2001; 2001US-00829472.
09-APR-2001; 2001US-0282698P.

04-MAY-2001; 2001US-0288590P.
29-MAY-2001; 2001US-0294443P.

(EOSB-) EOS BIOTECHNOLOGY INC.
Mack DH, Gish KC, Afar D;

WPI; 2002-583738/62.
N-PSDB; ABT07755.

Detecting a breast cancer-associated transcript in a patient's cell,
useful for diagnosing breast cancer, comprises contacting a biological

sample with a polynucleotide that selectively hybridizes with breast
cancer nucleic acids.

Disclosure; Page 405; 414pp; English.

The invention comprises a method of detecting a breast cancer-associated
transcript in a cell from a patient. The method of the invention involves

contacting a biological sample from the patient with a nucleotide that
hybridizes to one of the 69 breast cancer-associated gene sequences shown

in the specification. The method of the invention is useful in the
diagnosis or prognosis of breast cancer, and for detecting genes that are

up or down-regulated in breast cancer cells. Genes identified by the
method of the invention can be used in diagnostic purposes and also as

targets for screening for therapeutic compounds that modulate breast
cancer (e.g. hormones or antibodies). Identification of genes that are

over or under expressed in breast cancer can additionally provide high-
resolution, high-sensitivity datasets which can be used in the areas of

diagnostics, therapeutics, drug development, pharmacogenetics, protein
structure and biosensor development. Amino acid sequences ABJ05536 -

ABJ05604 represent the proteins encoded by the 69 breast cancer-
associated genes of the invention

Sequence 829 AA;

Query Match 100.0%; Score 4369; DB 5; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MGLPRGLASLLLVLCVLCQAASPCRAVFEAEVTLFAGGAEQPGQALGVFGCPG 60

1 MGLPRGLASLLLVLCVLCQAASPCRAVFEAEVTLFAGGAEQPGQALGVFGCPG 60

61 QEPALFSTNDNDDFTVRNGETVQERRSLKERNPLKIFPPSKRILRRHKRDVWVAPISVPENG 120

61 QEPALFSTNDNDDFTVRNGETVQERRSLKERNPLKIFPPSKRILRRHKRDVWVAPISVPENG 120

121 KGPPFORLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLINLKLPLDREI 180

121 KGPPFORLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLINLKLPLDREI 180

181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSLVGLPGTSGVMQV 240

181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSLVGLPGTSGVMQV 240

241 DEDDAIYTYNGVAVYSIHQEPKDPHDLMTFTHRSTGTISVSSGLDREKVPYETL 300

241 DEDDAIYTYNGVAVYSIHQEPKDPHDLMTFTHRSTGTISVSSGLDREKVPYETL 300

301 TMMDGSGSTTTAVAVVEILDANDNAPMFPDQKYEAHVPENAVGVHVEVQRLTV 360

301 TMMDGSGSTTTAVAVVEILDANDNAPMFPDQKYEAHVPENAVGVHVEVQRLTV 360

361 AWRATYLMGGDDGDHFTTTHPESNOGILTRKGLDFAKQHTLYVEVTNEAPFV 420

361 AWRATYLMGGDDGDHFTTTHPESNOGILTRKGLDFAKQHTLYVEVTNEAPFV 420

[illegible]

CC	solid surface with human and nonhuman cells. The invention is also useful
CC	for achieving better utilisation of vascular grafts and also in any
CC	system that employs endothelial cell-coated surface susceptible to
CC	hydrodynamic shear, such as shaking container coated with endothelial
CC	cells, continuous harvest systems for harvesting the products of
CC	endothelial cells on solid surface, and laminar shear devices. The
CC	present amino acid sequence represents the human P-cadherin (placental
XX	cadherin) protein of the invention
SQ	Sequence 829 AA;
	Query Match. 100.0%; Score 4369; DB 5; Length 829;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MGPRGPLASLLLOVCWLCQAASEPCRAVFAEAEVTLEAGGAEOEPGQALGVKFMCGPG 60
DB	1 MGPRGPLASLLLOVCWLCQAASEPCRAVFAEAEVTLEAGGAEOEPGQALGVKFMCGPG 60
QY	61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPFSKRILRRHKRDMVVAPISVPENG 120
DB	61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPFSKRILRRHKRDMVVAPISVPENG 120
QY	121 KGPPQRLNOIKSNKDSDTKIFYSITGPGADSPPEGVFAVEKETGTMLLNKPLDREIAK 180
DB	121 KGPPQRLNLQKSKNDRDTKIFYSITGPGADSPPEGVFAVEKETGTMLLNKPLDREIAK 180
QY	181 YELFGHAVSENGASVEDPMNISIIIVTDONDHKKFTODTPRGSVLEGLVPGTSMQVTTAT 240
DB	181 YELFGHAVSENGASVEDPMNISIIIVTDONDHKKFTODTPRGSVLEGLVPGTSMQVTTAT 240
QY	241 DEDDAIYYNGVWAYSIHSGPKDPHDLMTFIHRSTGTISVISGLDREKVPEYTLTIQA 300
DB	241 DEDDAIYYNGVWAYSIHSGPKDPHDLMTFIHRSTGTISVISGLDREKVPEYTLTIQA 300
QY	301 TDMDCGSTTTAVAVEIILDANDNAPMFDPOKYEAHPENAVGHVEQRILTVDLDA NSP 360
DB	301 TDMDCGSTTTAVAVEIILDANDNAPMFDPOKYEAHPENAVGHVEQRILTVDLDA NSP 360
QY	361 AWRATYLIMGDDGDGHFTITHPESNGOILTRKGLDFEAKNOHTLVXEVNTNEAPFVLKL 420
DB	361 AWRATYLIMGDDGDGHFTITHPESNGOILTRKGLDFEAKNQHTLVXEVNTNEAPFVLKL 420
QY	421 PTSTATIVHVHVEDVNEAPFVPPSKVVEQGEGPTGPFCVYTAAEDPDKENQKISYRILR 480
DB	421 PTSTATIVHVHVEDVNEAPFVPPSKVVEQGEGPTGPFCVYTAAEDPDKENQKISYRILR 480
QY	481 DPAGWLAMPDSGGTVATGILDRDEGOFVRNNIYEVNVLAMNDGSPPTTGTTLLLTLD 540
DB	481 DPAGWLAMPDSGGTVATGILDRDEGOFVRNNIYEVNVLAMNDGSPPTTGTTLLLTLD 540
QY	541 VNDHGVPPEPROITICNQSPVRHVNLITDKOLSPHTSFFOAQLTDDSDIYWTAENVBEGD 600
DB	541 VNDHGVPPEPROITICNQSPVRHVNLITDKOLSPHTSFFOAQLTDDSDIYWTAENVBEGD 600
QY	601 TVVLSSKKFLKQDITYDVHLSSLSDHGNKEQLTVIRATVCDCHGHVETCPGPWKGFILPV L 660
DB	601 TVVLSSKKFLKQDITYDVHLSSLSDHGNKEQLTVIRATVCDCHGHVETCPGPWKGFILPV L 660
QY	661 GAVLALLFLLLVLVLLLYRKRRKIKEPIILLPEDDPRDNVIFYGEGGGEEDDYDITOLHR 720
DB	661 GAVLALLFLLLVLVLLLYRKRRKIKEPIILLPEDDPRDNVIFYGEGGGEEDDYDITOLHR 720
QY	721 GLEARPEVLNRNDVAPTIIPTPMYRP RPANPDEIGNFTIENLKAANTDPTAPPYDTLLVF 780
DB	721 GLEARPEVLNRNDVAPTIIPTPMYRP RPANPDEIGNFTIENLKAANTDPTAPPYDTLLVF 780
QY	781 DYEGSGDAASSLSLTSSAQDQODYLYLNEWGRFKKLADMYGGGEDD 829
DB	781 DYEGSGDAASSLSLTSSAQDQODYLYLNEWGRFKKLADMYGGGEDD 829

XX 31-MAY-2001; 2001US-0294225P.
XX (CHIR) CHIRON CORP.
XX Reinhard C, Klinger J, Jefferson AB, Escobedo J, Randazo F;
XX Winter J, Goodson R;
XX WPI; 2003-140501/13.
XX N-PSDB; ABZ24736.
XX Inhibiting migration and proliferation of P-cadherin expressing cancer
XX for treating cancer, especially digestive cancer, characterized by
XX overexpression of P-cadherin, involves administering a P-cadherin
XX antagonist.
XX Disclosure; Page 10-11; 129pp; English.
XX The present sequence is the protein sequence for human P-cadherin
XX (placental cadherin), a calcium-dependent cellular adhesion protein. The
XX invention provides methods of treating or diagnosing cancers involving P-
XX cadherin expression using ligands that target P-cadherin, especially
XX human anti-P-cadherin antibodies. A claimed method of treating a cancer
XX characterised by the overexpression and/or upregulation of P-cadherin
XX comprises the administration of a P-cadherin antagonist, optionally
XX conjugated to a therapeutic agent. The migration, adhesion and/or
XX proliferation of the cancer is inhibited, and the method is especially
XX useful for treating or preventing a digestive cancer such as colon or
XX colorectal cancer. The antagonists may be an anti-P-cadherin antibody or
XX its fragment, a ribozyme or antisense oligonucleotide. A transgenic
XX animal that expresses a recombinant (human) antibody that specifically
XX binds P-cadherin is also claimed. A method of determining the presence of
XX cancer involves determining the expression level of P-cadherin in a cell
XX sample. Also provided are screens for identifying anti-P-cadherin
XX antibodies have therapeutic activity
XX Sequence 829 AA;

Query Match 100.0%; Score 4369; DB 6; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLPRGPLASLLQLVCWLOCAAGPCRAVFAEAVTLEAGGAEQEPQALGKVFMC 60
DB 1 MGLPRGPLASLLQLVCWLOCAAGPCRAVFAEAVTLEAGGAEQEPQALGKVFMC 60
QY 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG 120
DB 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG 120
QY 121 KGPPFQRLNQLKSKNRDRTKIFYGITGPGADSPPEGVFAVEKTCGWLNLNKPDLREBIAK 180
DB 121 KGPPFQRLNQLKSKNRDRTKIFYGITGPGADSPPEGVFAVEKTCGWLNLNKPDLREBIAK 180
QY 181 YELFGHVAENGASVEDPMNISITVTDQNDHKPKFTQDTPRGSLVGLPGTSVMQVAT 240
DB 181 YELFGHVAENGASVEDPMNISITVTDQNDHKPKFTQDTPRGSLVGLPGTSVMQVAT 240
QY 241 DEDDAITYNGVAVSIHQSPKDPKPHDLMTFTHIRSTGTISVISSGLDREKVPYTLTQA 300
DB 241 DEDDAITYNGVAVSIHQSPKDPKPHDLMTFTHIRSTGTISVISSGLDREKVPYTLTQA 300
QY 301 TDMGDGSGTTAVAVVEILDANDNAMPDPKQYEAHVPENAVGHVQRLTVDLDAENSP 360
DB 301 TDMGDGSGTTAVAVVEILDANDNAMPDPKQYEAHVPENAVGHVQRLTVDLDAENSP 360
QY 361 AWRATYLLMGDDGDHFTTTHPSNQGLITTRKGLDPEAKNQHTLYVEVNEAPFVLKL 420
DB 361 AWRATYLLMGDDGDHFTTTHPSNQGLITTRKGLDPEAKNQHTLYVEVNEAPFVLKL 420
QY 421 PTSTATIYVHVVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDDPKENQKISYRILR 480
DB 421 PTSTATIYVHVVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDDPKENQKISYRILR 480

QY 481 DPAGWLAMPDSDGQVTAAGTLDREDSQFVRNNIYEVWLAMDNGSPPTTGTGTLTLLID 540
DB 481 DPAGWLAMPDSDGQVTAAGTLDREDSQFVRNNIYEVWLAMDNGSPPTTGTGTLTLLID 540
QY 541 VNDHGFPVPEPRQITICNQSPVRHLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD 600
DB 541 VNDHGFPVPEPRQITICNQSPVRHLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD 600
QY 601 TVVLSLKKFLKQDYDVHLSLDHGKNEQLTVIRATVCDCHGHVETCPGPKGGFLLPVL 660
DB 601 TVVLSLKKFLKQDYDVHLSLDHGKNEQLTVIRATVCDCHGHVETCPGPKGGFLLPVL 660
QY 661 GAVLALLFLLLVLLLVLRKKIKPELLLPEDDTRDNVFFYEGEGEEDQDYDITQLHR 720
DB 661 GAVLALLFLLLVLLLVLRKKIKPELLLPEDDTRDNVFFYEGEGEEDQDYDITQLHR 720
QY 721 GLEARPEVVLNRDVAPTIIPMPYRPRANPDIGNFIENLKAANTDTPAPPYDTLLVF 780
DB 721 GLEARPEVVLNRDVAPTIIPMPYRPRANPDIGNFIENLKAANTDTPAPPYDTLLVF 780
QY 781 DYEGSGSDAASLSLSSASDQDDYDYLNEWGSRPKKLADMYGGGEDD 829
DB 781 DYEGSGSDAASLSLSSASDQDDYDYLNEWGSRPKKLADMYGGGEDD 829
RESULT 7
ADC15497
ID ADC15497 standard; protein; 829 AA.
XX
AC ADC15497;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human basal cell marker Cadherin-3.
XX
KW Human; basal cell marker; cadherin-3; tumour; P-cadherin;
KW matrix metalloproteinase 14;
KW cadherin EGF LAG seven-pass G-type receptor 2;
KW EGF-like domain Multiple 2; mammary lactation gland; breast tumour;
KW cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN US2003086934-A1.
XX
PD 08-MAY-2003.
XX
PF 26-JUL-2001; 2001US-00916849.
XX
PR 26-JUL-2000; 2000US-0220967P.
XX
PA (BOTS/) BOTSTEIN D.
PA (BROW/) BROWN P O.
PA (PERO/) PEROU C M.
PA (RING/) RING B.
PA (ROSS/) ROSS D.
PA (SEIT/) SEITZ R.
PA (VRIJ/) VAN DE RIJN J M.
XX
PI Botstein D, Brown PO, Perou CM, Ring B, Ross D, Seitz R;
PI Van De Rijn JM;
XX
XX WPI; 2003-786931/74.
XX
XX Classifying breast tumor by detecting expression or activity of gene
XX encoding cadherin 3, matrix metalloproteinase 14, cadherin EGF LAG seven-
XX pass G-type receptor 2 in tumor sample.
XX
XX Claim 1; SEQ ID NO 1; 88pp; English.
XX
XX The invention relates to classifying (M1) a tumour comprising detecting
XX expression or activity of a gene encoding cadherin 3 or P-cadherin
CC

CC ,matrix metalloproteinase 14 , or cadherin EGF LAG seven-pass G-type
CC receptor 2 or EGF-like domain, Multiple 2 (P1-3) or at least two genes
CC chosen from gene encoding (P1), (P2) or (P3) in the tumour sample, and
CC classifying tumour as belonging to a tumour subclass based on results of
CC detecting step. Also included are testing a subject for expression of P1-
CC P3 and providing diagnostic, prognostic, or predictive information based
CC on the detecting step, stratifying the subject for a clinical trial based
CC on the detecting step, or selecting a treatment based on the detecting
CC step, an antibody that specifically binds to an epitope found in (P1),
CC (P2) or (P3) (where the antibody recognises basal cells in normal mammary
CC lactation glands), methods of testing a compound or a combination of
CC compounds for activity against tumours. The method (M1) is useful for
CC classifying a tumour (breast tumour), as belonging to a tumour subclass
CC such as basal tumour subclass. The antibody is useful for treating a
CC subject in need of treatment for cancer. The tumour is a breast tumour,
CC and where the method further involves identifying the tumour as belonging
CC to a basal tumour subclass. The antibody is conjugated with a toxic
CC molecule. The present sequence is human Cadherin-3.
XX
SQ Sequence 829 AA;
Query Match 100.0%; Score 4369; DB 7; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLPRGLASLLQLQVLCQAASEPCRAVFREAEVTLFAGGAEQPGQALGKVFNGCPG 60
Db 1 MGLPRGLASLLQLQVLCQAASEPCRAVFREAEVTLFAGGAEQPGQALGKVFNGCPG 60
QY 61 QEPALFSTNDNDFTVRNGETVQRRSLKERNPLKIPFSKRILRRHKRDVWVAPISVPENG 120
Db 61 QEPALFSTNDNDFTVRNGETVQRRSLKERNPLKIPFSKRILRRHKRDVWVAPISVPENG 120
QY 121 KGPFPORLNQKSKNDRTKIFYSITGPGADSPGEGVFAVEKETGWLKLNKPLDREIEIAK 180
Db 121 KGPFPORLNQKSKNDRTKIFYSITGPGADSPGEGVFAVEKETGWLKLNKPLDREIEIAK 180
QY 181 YELFHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTRFGSVLFGVLPSTVMQVTAT 240
Db 181 YELFHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTRFGSVLFGVLPSTVMQVTAT 240
QY 241 DEDDAITYNGVAVYSHSQEPKDPHDLMFTIHRSTGTISVSSGLDRKVEPYTLTIOA 300
Db 241 DEDDAITYNGVAVYSHSQEPKDPHDLMFTIHRSTGTISVSSGLDRKVEPYTLTIOA 300
QY 301 TMDGPGSTTTAVAVVEILDANDNAPMDPQKYEAVHPENAVGHEVQRLTVTDLDPNSP 360
Db 301 TMDGPGSTTTAVAVVEILDANDNAPMDPQKYEAVHPENAVGHEVQRLTVTDLDPNSP 360
QY 361 AWRATYLMGGDDGDHFTITTHPESNQGLITTRKGLDFAKNGHTLYVEVTNEAPFVLKL 420
Db 361 AWRATYLMGGDDGDHFTITTHPESNQGLITTRKGLDFAKNGHTLYVEVTNEAPFVLKL 420
QY 421 PTSTATIVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEPDPEKNIKISVRIIR 480
Db 421 PTSTATIVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEPDPEKNIKISVRIIR 480
QY 481 DPAGWLAMPDSCGVTAAGTLDREDSQFVRNNIYEVWVLAMONGSPPTTGTGTLTLTLD 540
Db 481 DPAGWLAMPDSCGVTAAGTLDREDSQFVRNNIYEVWVLAMONGSPPTTGTGTLTLTLD 540
QY 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDSDIYWTAEVNEEGD 600
QY 601 TVVLSLKFKLQDITYDHLSSLSDHGNKEQLTVIRATVCDCHGVETCPGPKWGGFLLPVL 660
Db 601 TVVLSLKFKLQDITYDHLSSLSDHGNKEQLTVIRATVCDCHGVETCPGPKWGGFLLPVL 660
QY 661 GAVLALLFLLVLLVLLVLRKKRKEPLLPEDDTRDNVIFYEGEGGEEDQDYDTQLHR 720
Db 661 GAVLALLFLLVLLVLLVLRKKRKEPLLPEDDTRDNVIFYEGEGGEEDQDYDTQLHR 720

QY 721 GLEARPEVVLNRNDVAPTIPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVVLNRNDVAPTIPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
QY 781 DYEGSGSDAASLSLTSASDQDDYDVLNWSRFRKFLADMYGGGEDD 829
Db 781 DYEGSGSDAASLSLTSASDQDDYDVLNWSRFRKFLADMYGGGEDD 829
RESULT 8
ADD14190
ID ADD14190 standard; protein; 829 AA.
XX
AC ADD14190;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human src biomarker polypeptide SEQ ID NO:379.
XX
KW predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human.
XX
OS Homo sapiens.
XX
PN WO2003062395-A2.
XX
PD 31-JUL-2003.
XX
PF 17-JAN-2003; 2003WO-US001981.
XX
PR 18-JAN-2002; 2002US-0350061P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Huang F, Fairchild CR, Lee FY, Shaw P;
XX
WI WPI; 2003-636735/60.
DR N-PSDB; ADD14795.
XX
PT New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.
XX
PS Claim 10; SEQ ID NO 379; 139pp; English.
XX
CC The present invention describes a predictor set comprising a plurality of
CC polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.

QY	1	MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLBAGAEQPGQALGKVFMGCPG	60
Db	1	MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLBAGAEQPGQALGKVFMGCPG	60
QY	61	QBPALFSTNDNDFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDWVAPISVPENG	120
Db	61	QBPALFSTNDNDFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDWVAPISVPENG	120
QY	121	KGPPFORLNQKSNKDRDKIFYSITGPGADSPPEGFAVEKETGWLILNKKPLDREEIAK	180
Db	121	KGPPFORLNQKSNKDRDKIFYSITGPGADSPPEGFAVEKETGWLILNKKPLDREEIAK	180
QY	181	YELFGHAVSENGASVEDPNNISIIIVTDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVTTAT	240
Db	181	YELFGHAVSENGASVEDPNNISIIIVTDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVTTAT	240
QY	241	DEDDALTYTNGVVAVYSHSQEPKDPDHLMTFTHRSSTGTSVSSGIDREKVEYTLTIOA	300
Db	241	DEDDALTYTNGVVAVYSHSQEPKDPDHLMTFTHRSSTGTSVSSGIDREKVEYTLTIOA	300
QY	301	TMDGSGSTTTAVAVVEILLDANDNPMFDPQKYEAHVPENAVGHEVQRLTVTDLAPNSP	360
Db	301	TMDGSGSTTTAVAVVEILLDANDNPMFDPQKYEAHVPENAVGHEVQRLTVTDLAPNSP	360
QY	361	AWRATYLIMGDDGDHFTTTHPESNQGLITRKGDLDFEAKNQHTLYVEVTNEAPFVLKL	420
Db	361	AWRATYLIMGDDGDHFTTTHPESNQGLITRKGDLDFEAKNQHTLYVEVTNEAPFVLKL	420
QY	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGTPTGEPVCVVTAEADPKENKISRIILR	480
Db	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGTPTGEPVCVVTAEADPKENKISRIILR	480
QY	481	DPAGWLAMPDQSGQVAVGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGLTLTLID	540
Db	481	DPAGWLAMPDQSGQVAVGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGLTLTLID	540
QY	541	VNDHGVPPEPRQITICNQSPVRHVLNITDKOLSPHTSPFQAQITDSDIYWTAEVNEEGD	600
Db	541	VNDHGVPPEPRQITICNQSPVRHVLNITDKOLSPHTSPFQAQITDSDIYWTAEVNEEGD	600
QY	601	TVVLSLKKFLKQDTYDVHLSLDHGNKEQLTVIRATVCDCHGVETCPGPGKGFILPVL	660
Db	601	TVVLSLKKFLKQDTYDVHLSLDHGNKEQLTVIRATVCDCHGVETCPGPGKGFILPVL	660
QY	661	GAVALLLFLLVLLVLRKKRIKEPLLPEDDTRDNVFFYEGEGGEDQDYDITQLHR	720
Db	661	GAVALLLFLLVLLVLRKKRIKEPLLPEDDTRDNVFFYEGEGGEDQDYDITQLHR	720
QY	721	GLEARPEVVLNDVAPTIITPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF	780
Db	721	GLEARPEVVLNDVAPTIITPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF	780
QY	781	DYEGSGSDAASLSLTSASDQDDYDYLNEGSRFKKLADMYGGGEDD	829
Db	781	DYEGSGSDAASLSLTSASDQDDYDYLNEGSRFKKLADMYGGGEDD	829
RESULT 10			
ID	ADN39578	standard; protein; 829 AA.	
AC	AC	ADN39578;	
DT	DT	17-JUN-2004 (first entry)	
XX	XX	Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A178.	
XX	XX	Human; differential expression; cancer; angiogenic disorder;	
KW	KW	fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;	
KW	KW	inflammatory disease; autoimmune disease;	
KW	KW	retinal neovascularisation syndrome; scarring; uterine fibroid;	
KW	KW	detection; diagnosis; prognosis; drug screening; drug targeting;	

KW	wound healing; contraception; cytostatic; cardiant; immunomodulatory;		
KW	vulnerary; gene therapy; vaccine.		
OS	Homo sapiens.		
XX	WO2003042661-A2.		
PN	22-MAY-2003.		
XX	13-NOV-2002; 2002WO-US036810.		
XX	13-NOV-2001; 2001US-0350666P.		
PR	21-NOV-2001; 2001US-0332464P.		
PR	29-NOV-2001; 2001US-0334393P.		
PR	03-DEC-2001; 2001US-0335394P.		
PR	14-DEC-2001; 2001US-0340376P.		
PR	08-JAN-2002; 2002US-0347211P.		
PR	10-JAN-2002; 2002US-0347349P.		
PR	08-FEB-2002; 2002US-0355250P.		
PR	13-FEB-2002; 2002US-0356714P.		
PR	20-FEB-2002; 2002US-0359077P.		
PR	29-MAR-2002; 2002US-0368099P.		
PR	04-APR-2002; 2002US-0370110P.		
PR	12-APR-2002; 2002US-0372246P.		
PR	05-JUN-2002; 2002US-0386614P.		
PR	16-JUL-2002; 2002US-0396839P.		
PR	22-JUL-2002; 2002US-039775P.		
PR	22-JUL-2002; 2002US-0397845P.		
PR	09-SEP-2002; 2002US-0409450P.		
XX	(EOSB-) EOS BIOTECHNOLOGY INC.		
PA	Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;		
PI	Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;		
XX	WPI: 2003-468649/44.		
DR	N-PSDB; ADN39577.		
XX	Determining the presence or absence of a pathological cell in a patient,		
PT	useful for diagnosing, prognosing or treating cancer, comprises detecting		
PT	a nucleic acid in a biological sample.		
XX	Claim 12; SEQ ID NO A178; 1385pp; English.		
XX	The invention relates to nucleic acids and proteins (ADN39683-ADN40064)		
CC	whose expression is upregulated or downregulated in specific cancers or		
CC	other diseases such as angiogenic or fibrotic disorders, and to methods		
CC	of determining the presence or absence of a pathological cell in a		
CC	patient by detecting a nucleic acid at least 80% identical to those of		
CC	the invention or by detecting a polypeptide of the invention. The		
CC	invention also relates to expression vectors and host cells comprising a		
CC	nucleic acid of the invention; antibodies which specifically bind a		
CC	polypeptide of the invention; use of such antibodies for drug targeting;		
CC	and methods of screening for modulators of activity or expression of the		
CC	polypeptides and nucleic acids. The nucleic acids, polypeptides,		
CC	antibodies and methods are useful for diagnosing, prognosing and treating		
CC	cancer and other conditions such as psoriasis, ischaemia, heart disease,		
CC	atherosclerosis, inflammatory diseases, autoimmune diseases, retinal		
CC	neovascularisation syndromes, scarring and uterine fibroids. They may		
CC	also be useful in wound healing and in contraception. The present		
CC	sequence represents a polypeptide of the invention.		
XX	Sequence 829 AA;		
QY	Query Match 100.0%; Score 4369; DB 7; Length 829;		
Db	Best Local Similarity 100.0%; Pred. No. 0;		
QY	Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLBAGAEQPGQALGKVFMGCPG	60
Db	1	MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLBAGAEQPGQALGKVFMGCPG	60
QY	61	QBPALFSTNDNDFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDWVAPISVPENG	120

```
Db 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRLRRHKRDVWVAPISVPENG 120
Qy 121 KGPPQRLNOLKSNKDRDTKIFYSITGPGADSPGEGVFAVEKETGWLKLLNKPDLREETAK 180
Db 121 KGPPQRLNOLKSNKDRDTKIFYSITGPGADSPGEGVFAVEKETGWLKLLNKPDLREETAK 180
Qy 181 YELFGHAVSENGAVEDEPMNISIIIVTQNDHKPKFTQDTERGSLVLEGVLPGTSMQVAT 240
Db 181 YELFGHAVSENGAVEDEPMNISIIIVTQNDHKPKFTQDTERGSLVLEGVLPGTSMQVAT 240
Qy 241 DEDDAIYTYNGVWAYSIIHQEPKDPHLMFTIHRSTGTISVSSGLDREKVPYTLTIA 300
Db 241 DEDDAIYTYNGVWAYSIIHQEPKDPHLMFTIHRSTGTISVSSGLDREKVPYTLTIA 300
Qy 301 TDMDCGSGTTTAVAVVEILDANDNAPFDPQKYEAHPVNAVGHVEQRLTWTDLDAENSP 360
Db 301 TDMDCGSGTTTAVAVVEILDANDNAPFDPQKYEAHPVNAVGHVEQRLTWTDLDAENSP 360
Qy 361 AWRATYILMGDDGDHFTITTHPSNQGILITRKGLOFEAKNQHTLVVETNEAPFVLKL 420
Db 361 AWRATYILMGDDGDHFTITTHPSNQGILITRKGLOFEAKNQHTLVVETNEAPFVLKL 420
Qy 421 PTSTATIVHVEDVNEAPVFVPPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480
Db 421 PTSTATIVHVEDVNEAPVFVPPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480
Qy 481 DPAGWLAMPDPSGQVTAAGTLDREDEQFVRNNIYEVNVLAMNDGSPPTGTGTLTLLID 540
Db 481 DPAGWLAMPDPSGQVTAAGTLDREDEQFVRNNIYEVNVLAMNDGSPPTGTGTLTLLID 540
Qy 541 VNDHGPVPEPQIITICNQSPPRVHVNITDKLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPQIITICNQSPPRVHVNITDKLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSLKKFLKQDTYDVHLSLSDRHNEQLTVIRATVCDCHGHVETCPGPKWGFIPLVL 660
Db 601 TVVLSLKKFLKQDTYDVHLSLSDRHNEQLTVIRATVCDCHGHVETCPGPKWGFIPLVL 660
Qy 661 GAVLALLFLVLLVLLVLRKXKIKEPILLPEDDTRDNVYVYGGEGGEEDDYITOLHR 720
Db 661 GAVLALLFLVLLVLLVLRKXKIKEPILLPEDDTRDNVYVYGGEGGEEDDYITOLHR 720
Qy 721 GLARPEVLRNDVAPTIPTPMYRPRPANPDEIGNFTIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLARPEVLRNDVAPTIPTPMYRPRPANPDEIGNFTIENLKAANTDPTAPPYDTLLVF 780
Qy 781 DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSFRFKKLADMYCGGEDD 829
Db 781 DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSFRFKKLADMYCGGEDD 829

RESULT 11
ADN39545
ID ADN39545 standard; protein; 829 AA.
XX AC ADN39545;
XX DT 17-JUN-2004 (first entry)
XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A145.
XX KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnary; gene therapy; vaccine.
XX OS Homo sapiens.
XX PN WO2003042661-A2.
```

```
XX 22-MAY-2003.
XX 13-NOV-2002; 2002WO-US036810.
XX 13-NOV-2001; 2001US-0350666P.
XX 21-NOV-2001; 2001US-0332464P.
XX 29-NOV-2001; 2001US-0334393P.
XX 03-DEC-2001; 2001US-0335394P.
XX 14-DEC-2001; 2001US-0340376P.
XX 08-JAN-2002; 2002US-0347211P.
XX 10-JAN-2002; 2002US-0347349P.
XX 08-FEB-2002; 2002US-035250P.
XX 13-FEB-2002; 2002US-0356714P.
XX 20-FEB-2002; 2002US-0359077P.
XX 29-MAR-2002; 2002US-0368809P.
XX 04-APR-2002; 2002US-0370110P.
XX 12-APR-2002; 2002US-0372246P.
XX 05-JUN-2002; 2002US-0386614P.
XX 16-JUL-2002; 2002US-0396839P.
XX 22-JUL-2002; 2002US-0397775P.
XX 22-JUL-2002; 2002US-0397845P.
XX 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
XX Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX WPI; 2003-468649/44.
XX N-PSDB; ADN39544.
XX Determining the presence or absence of a pathological cell in a patient,
XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.
XX Claim 12; SEQ ID NO A145; 1385pp; English.
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
XX whose expression is upregulated or downregulated in specific cancers or
XX other diseases such as angiogenic or fibrotic disorders, and to methods
XX of determining the presence or absence of a pathological cell in a
XX patient by detecting a nucleic acid at least 80% identical to those of
XX the invention or by detecting a polypeptide of the invention. The
XX invention also relates to expression vectors and host cells comprising a
XX nucleic acid of the invention; antibodies which specifically bind a
XX polypeptide of the invention; use of such antibodies for drug targeting;
XX and methods of screening for modulators of activity or expression of the
XX polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX antibodies and methods are useful for diagnosing, prognosing and treating
XX cancer and other conditions such as psoriasis, ischaemia, heart disease,
XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX neovascularisation syndromes, scarring and uterine fibroids. They may
XX also be useful in wound healing and in contraception. The present
XX sequence represents a polypeptide of the invention.
XX Sequence 829 AA;
XX Query Match 100.0%; Score 4369; DB 7; Length 829;
XX Best Local Similarity 100.0%; Pred. NO. 0;
XX Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGUPRGPLASILLQVCWLOCAASEPCRAVFREAEVTEAGGAEQEPQALGKVFMCPCG 60
Db 1 MGUPRGPLASILLQVCWLOCAASEPCRAVFREAEVTEAGGAEQEPQALGKVFMCPCG 60
Qy 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRLRRHKRDVWVAPISVPENG 120
Db 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRLRRHKRDVWVAPISVPENG 120
Qy 121 KGPPQRLNOLKSNKDRDTKIFYSITGPGADSPGEGVFAVEKETGWLKLLNKPDLREETAK 180
Db 121 KGPPQRLNOLKSNKDRDTKIFYSITGPGADSPGEGVFAVEKETGWLKLLNKPDLREETAK 180
```


QY 181 YELFGHAVSENGASVEDPMNISIIVTDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVTTAT 240
DB 181 YELFGHAVSENGASVEDPMNISIIVTDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVTTAT 240
QY 241 DEDDALTYNGVAVYSHSQEPKDPHDLMTFTHRSSTGTSVSSGDLREKVPETLTIOA 300
DB 241 DEDDALTYNGVAVYSHSQEPKDPHDLMTFTHRSSTGTSVSSGDLREKVPETLTIOA 300
QY 301 TMDGSGSTTTAVAVVEILDANDNPMFDPQKYEAHVPENAVGHVQRLTVTDLDAVNSP 360
DB 301 TMDGSGSTTTAVAVVEILDANDNPMFDPQKYEAHVPENAVGHVQRLTVTDLDAVNSP 360
QY 361 AWRATYLINGDDGHTTTTHPESQGLTTRKGLDFAKQHTLYVEVTNEAPVPLKL 420
DB 361 AWRATYLINGDDGHTTTTHPESQGLTTRKGLDFAKQHTLYVEVTNEAPVPLKL 420
QY 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGPTGEPVCVVTAEPPDKENOKISVRIILR 480
DB 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGPTGEPVCVVTAEPPDKENOKISVRIILR 480
QY 481 DPAGWLAMPDSCQVTAAGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLLTLD 540
DB 481 DPAGWLAMPDSCQVTAAGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLLTLD 540
QY 541 VNDHGVPPEPRQITICNQSPPVRHVLNITDKDLSPTSPPFQAQLTDDSDIYWTAEVNEEGD 600
DB 541 VNDHGVPPEPRQITICNQSPPVRHVLNITDKDLSPTSPPFQAQLTDDSDIYWTAEVNEEGD 600
QY 601 TVVLSLKKFLKQDVTYVHLSLSDHGKNEQLTVIRATVCDCHGVETCPGPKGGFLPLVL 660
DB 601 TVVLSLKKFLKQDVTYVHLSLSDHGKNEQLTVIRATVCDCHGVETCPGPKGGFLPLVL 660
QY 661 GAVLALLFLLVLLVLRKKRKEKPELLPDDTRDNVFFYBEGGEGEDQYDITQLHR 720
DB 661 GAVLALLFLLVLLVLRKKRKEKPELLPDDTRDNVFFYBEGGEGEDQYDITQLHR 720
QY 721 GLEARPEVVLNDVAPTIPTMYRPRPANPBEIGNFIENLKAANTDPTAPPYDPLLVF 780
DB 721 GLEARPEVVLNDVAPTIPTMYRPRPANPBEIGNFIENLKAANTDPTAPPYDPLLVF 780
QY 781 DYEGSGSDAASLSLSTSSASDQDDQDYDLYNWSRFPKLLADMYGGGEDD 829
DB 781 DYEGSGSDAASLSLSTSSASDQDDQDYDLYNWSRFPKLLADMYGGGEDD 829
RESULT 12
ADN39465
AC ADN39465 standard; protein; 829 AA.
XX
XX ADN39465;
DT 17-JUN-2004 (first entry)
XX
XX Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A65.
DE
DE Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
KW vulnery; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
XX WO2003042661-A2.
XX
XX 22-MAY-2003.
XX
XX 13-NOV-2002; 2002WO-US036810.
XX
XX 13-NOV-2001; 2001US-0350666P.

PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-035250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-039775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
PA Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI: 2003-468649/44.
DR N-PSDB; ADN39464.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 12; SEQ ID NO A65; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN39683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides, and
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX
SQ Sequence 829 AA;
Query Match 100.0%; Score 4369; DB 7; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLPRGLASLALLLQVCWLQCAASPCRAVFEAEVTLAEAGGAEQPGQALGKVPFGCPG 60
DB 1 MGLPRGLASLALLLQVCWLQCAASPCRAVFEAEVTLAEAGGAEQPGQALGKVPFGCPG 60
QY 61 QEPALFSTNDNDFTVRNGETVOERSLKERNPLKIPPSKRIILRRHKRDVWVAPISVPENG 120
DB 61 QEPALFSTNDNDFTVRNGETVOERSLKERNPLKIPPSKRIILRRHKRDVWVAPISVPENG 120
QY 121 KGPFPQRLNOLKSNKDRDTKIFYSITGCGADSPPGVFAVEKETGWLILINKLPDREEIAK 180
DB 121 KGPFPQRLNOLKSNKDRDTKIFYSITGCGADSPPGVFAVEKETGWLILINKLPDREEIAK 180
QY 181 YELFGHAVSENGASVEDPMNISIIVTDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVTTAT 240
DB 181 YELFGHAVSENGASVEDPMNISIIVTDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVTTAT 240
QY 241 DEDDALTYNGVAVYSHSQEPKDPHDLMTFTHRSSTGTSVSSGDLREKVPETLTIOA 300

Db 241 DEDDAIYVNGVWAYSISHSQEPKDPHDLMTFIHRSTGTISVISGLDREKVPVETLTIQA 300
Qy 301 TDMGDSGTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVEQRLTVTDLDAENSP 360
Db 301 TDMGDSGTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVEQRLTVTDLDAENSP 360
Qy 361 AWRATYILMGDDGDHFTIITHPESNQGILTRKGLDPEAKNQHTLVVEVTNEAPFVLKL 420
Db 361 AWRATYILMGDDGDHFTIITHPESNQGILTRKGLDPEAKNQHTLVVEVTNEAPFVLKL 420
Qy 421 PTSTATIVVHVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDPKENQKISYRILR 480
Db 421 PTSTATIVVHVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDPKENQKISYRILR 480
Qy 481 DPAGWLAMPDPSGQVTAAGTLDREDEQFVRNNIYEVNVLAMNDGSPPTTGTGTLTLLID 540
Db 481 DPAGWLAMPDPSGQVTAAGTLDREDEQFVRNNIYEVNVLAMNDGSPPTTGTGTLTLLID 540
Qy 541 VNDHGPVPEPQIITCNQSPVRHVNLITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPQIITCNQSPVRHVNLITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSLKKFLKQDITYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Db 601 TVVLSLKKFLKQDITYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Qy 661 GAVLALLFLLLVLLLVLRKGRKIKEPILLPDDTRDNVFFYGGEGGEEDQDITQLHR 720
Db 661 GAVLALLFLLLVLLLVLRKGRKIKEPILLPDDTRDNVFFYGGEGGEEDQDITQLHR 720
Qy 721 GLEARPEVLNDVAPTIIPPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDITLLVF 780
Db 721 GLEARPEVLNDVAPTIIPPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDITLLVF 780
Qy 781 DYEAGSDAASLSSITSSASDQDQDYDYNLWGSFRFKKLADMYGGGDD 829
Db 781 DYEAGSDAASLSSITSSASDQDQDYDYNLWGSFRFKKLADMYGGGDD 829

RESULT 13

ADL06565
ID ADL06565 standard; protein; 829 AA.
XX AC ADL06565;
XX DT 20-MAY-2004 (first entry)
XX DE Human tumour-associated antigenic target (TAT) polypeptide #64.
XX KW Human; tumour-associated antigenic target; TAT; cell death; tumour;
XX KW cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO2004016225-A2.
XX PD 26-FEB-2004.
XX PF 19-AUG-2003; 2003WO-US025892.
XX PR 19-AUG-2002; 2002US-0404809P.
PR 21-AUG-2002; 2002US-0405645P.
PR 23-SEP-2002; 2002US-0413192P.
PR 15-OCT-2002; 2002US-0419008P.
PR 15-NOV-2002; 2002US-0426847P.
XX PR 02-JUL-2003; 2003US-0484959P.
PA (GETH) GENENTECH INC.
XX Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;
PI Spencer SD, Wu TD, Zhang Z;
XX

DR WPI; 2004-257144/24.
DR N-PSDB; ADL06487.
XX New antibody that binds to a tumor-associated antigenic target (TAT) polypeptide, useful for preparing a composition for diagnosing or treating cancer.
PT
PT
XX Claim 2; SEQ ID NO 145; 319pp; English.
PS
XX The present invention relates to the isolation of human tumour-associated antigenic target (TAT) polynucleotide and polypeptide sequences. Also disclosed is an antibody that binds to a TAT polypeptide. The antibody is a monoclonal antibody, an antibody fragment, a chimeric antibody or a humanised antibody. It is conjugated to a growth inhibitory agent. It is produced in bacteria or in CHO cells and induces death of a cell to which it binds. The antibody is useful for preparing a composition for diagnosing or treating tumours and cancer. The present sequence represents a human TAT polypeptide of the invention.
CC
SQ Sequence 829 AA;

Query Match 100.0%; Score 4369; DB 8; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGLPRGPLASLLQLQVCWLOCAASEPCRAVFREAEVTLLEAGGAEQEPQALGKVFMGCPG 60
Db 1 MGLPRGPLASLLQLQVCWLOCAASEPCRAVFREAEVTLLEAGGAEQEPQALGKVFMGCPG 60
Qy 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPSKRILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPSKRILRRHKRDWVAPISVPENG 120
Qy 121 KGFPQRLNQLKSKORDTKIFYSITGPGADSPGEGVFAVEKETGWLILNKPLDREIAK 180
Db 121 KGFPQRLNQLKSKORDTKIFYSITGPGADSPGEGVFAVEKETGWLILNKPLDREIAK 180
Qy 181 YELFGHAVSENGASVEDPMNISIIITDONDHPKFTQDTFRGSLGVLGTPGTSVMQVAT 240
Db 181 YELFGHAVSENGASVEDPMNISIIITDONDHPKFTQDTFRGSLGVLGTPGTSVMQVAT 240
Qy 241 DEDDAIYVNGVWAYSISHSQEPKDPHDLMTFIHRSTGTISVISGLDREKVPVETLTIQA 300
Db 241 DEDDAIYVNGVWAYSISHSQEPKDPHDLMTFIHRSTGTISVISGLDREKVPVETLTIQA 300
Qy 301 TDMGDSGTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVEQRLTVTDLDAENSP 360
Db 301 TDMGDSGTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVEQRLTVTDLDAENSP 360
Qy 361 AWRATYILMGDDGDHFTIITHPESNQGILTRKGLDPEAKNQHTLVVEVTNEAPFVLKL 420
Db 361 AWRATYILMGDDGDHFTIITHPESNQGILTRKGLDPEAKNQHTLVVEVTNEAPFVLKL 420
Qy 421 PTSTATIVVHVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDPKENQKISYRILR 480
Db 421 PTSTATIVVHVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDPKENQKISYRILR 480
Qy 481 DPAGWLAMPDPSGQVTAAGTLDREDEQFVRNNIYEVNVLAMNDGSPPTTGTGTLTLLID 540
Db 481 DPAGWLAMPDPSGQVTAAGTLDREDEQFVRNNIYEVNVLAMNDGSPPTTGTGTLTLLID 540
Qy 541 VNDHGPVPEPQIITCNQSPVRHVNLITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPQIITCNQSPVRHVNLITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSLKKFLKQDITYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Db 601 TVVLSLKKFLKQDITYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Qy 661 GAVLALLFLLLVLLLVLRKGRKIKEPILLPDDTRDNVFFYGGEGGEEDQDITQLHR 720
Db 661 GAVLALLFLLLVLLLVLRKGRKIKEPILLPDDTRDNVFFYGGEGGEEDQDITQLHR 720

```
QY 721 GLEARPEVLRNDVAPTIPTMYRPRRPNPDEIGNFIENLKAANTDPTAPPYDILLVF 780
DB 721 GLEARPEVLRNDVAPTIPTMYRPRRPNPDEIGNFIENLKAANTDPTAPPYDILLVF 780
QY 781 DYEGSGDAASLSLTSASDQDDYDYLNEWGSRPKKLADMYGGGDD 829
DB 781 DYEGSGDAASLSLTSASDQDDYDYLNEWGSRPKKLADMYGGGDD 829

RESULT 14
ADQ20015
ID ADQ20015 standard; protein; 829 AA.
XX AC ADQ20015;
XX DT 26-AUG-2004 (first entry)
XX DE Human soft tissue sarcoma-upregulated protein - SEQ ID 2835.
XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX OS Homo sapiens.
XX PN WO2004048938-A2.
XX PD 10-JUN-2004.
XX PF 26-NOV-2003; 2003WO-US038193.
XX PR 26-NOV-2002; 2002US-0429739P.
XX (PROT-) PROTEIN DESIGN LABS INC.
XX PA Aziz N, Ginsburg WM, Zlotnik A;
XX PI WPI; 2004-441208/41.
XX DR Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX Example 2; SEQ ID NO 2835; 210pp; English.
XX PS The invention relates to a novel method for detecting soft tissue sarcoma
XX CC which comprises obtaining a first soft tissue sample from an individual
XX CC and a normal soft tissue sample from the same or different individual,
XX CC determining the expression of a gene in both samples and comparing the
XX CC expression of the gene in both soft tissue samples, where a higher level
XX CC of protein expression in the first soft tissue sample indicates the
XX CC presence of soft tissue sarcoma. The method of the invention has
XX CC cytostatic applications and may be useful for detecting soft tissue
XX CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX CC acid sequences may be useful in diagnostic and screening applications.
XX CC The current sequence is that of a human soft tissue sarcoma-upregulated
XX CC protein of the invention. The current sequence is not shown within the
XX CC specification per se but was submitted in CD format by the inventor.
XX SQ Sequence 829 AA;
Query Match 100.0%; Score 4369; DB 8; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLPRGLASLLLVLCWLOCAASEPCRAVFRAEVTLEAGAEQSPGQALGVFMGCPG 60
DB 1 MGLPRGLASLLLVLCWLOCAASEPCRAVFRAEVTLEAGAEQSPGQALGVFMGCPG 60
QY 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKILRRHKRDWVAPISVPEG 120
DB 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKILRRHKRDWVAPISVPEG 120
QY 121 KGPPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDREBEIAK 180
```

```
DB 121 KGPPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDREBEIAK 180
QY 181 YELFGHAYSENGASVEDPMNISIIIVTDQNDHHPKFTQDTFRGSLVLEGLVLPGTSMQVTTAT 240
DB 181 YELFGHAYSENGASVEDPMNISIIIVTDQNDHHPKFTQDTFRGSLVLEGLVLPGTSMQVTTAT 240
QY 241 DEDDAIYTYNGVAVYIHSQEPKDPHDLMTTHRSTGTISVISSSGLDREKVPYETLTIOA 300
DB 241 DEDDAIYTYNGVAVYIHSQEPKDPHDLMTTHRSTGTISVISSSGLDREKVPYETLTIOA 300
QY 301 TDMGDSGTTTAVAVVEILDANDNAPMEDPQYEAHVPENAVGHVEVQRLTVTDLDPNSP 360
DB 301 TDMGDSGTTTAVAVVEILDANDNAPMEDPQYEAHVPENAVGHVEVQRLTVTDLDPNSP 360
QY 361 AWRATYLMGGDDGDHFTTTHPESNOGILLTTRKGLDFAKQKQHTLYVVTNEAPFVLKL 420
DB 361 AWRATYLMGGDDGDHFTTTHPESNOGILLTTRKGLDFAKQKQHTLYVVTNEAPFVLKL 420
QY 421 PTSTATIVVHVEDVNEAPVFPSPKVVEQEGIPGEPVCVYTAEDPDKENOKISYRIILR 480
DB 421 PTSTATIVVHVEDVNEAPVFPSPKVVEQEGIPGEPVCVYTAEDPDKENOKISYRIILR 480
QY 481 DPAGWLAMPDPSGQVTAVGTLDRDEQFVRNNIYEVWVLAMONGSPPTTGTGLTLTLID 540
DB 481 DPAGWLAMPDPSGQVTAVGTLDRDEQFVRNNIYEVWVLAMONGSPPTTGTGLTLTLID 540
QY 541 VNDHGPVPEPROITTCNQSPVRHVLNITDKOLSPHTSPFOAQLTDDSDIYWTAEVNEEGD 600
DB 541 VNDHGPVPEPROITTCNQSPVRHVLNITDKOLSPHTSPFOAQLTDDSDIYWTAEVNEEGD 600
QY 601 TVVLSLKFKLQDITYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFLLPVL 660
DB 601 TVVLSLKFKLQDITYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFLLPVL 660
QY 661 GAVLALLFLLLVLLVLRKKRKIKEPLLPEDDTRDNVFFYEGEGGEDDYDITQLHR 720
DB 661 GAVLALLFLLLVLLVLRKKRKIKEPLLPEDDTRDNVFFYEGEGGEDDYDITQLHR 720
QY 721 GLEARPEVLRNDVAPTIPTMYRPRRPNPDEIGNFIENLKAANTDPTAPPYDILLVF 780
DB 721 GLEARPEVLRNDVAPTIPTMYRPRRPNPDEIGNFIENLKAANTDPTAPPYDILLVF 780
QY 781 DYEGSGDAASLSLTSASDQDDYDYLNEWGSRPKKLADMYGGGDD 829
DB 781 DYEGSGDAASLSLTSASDQDDYDYLNEWGSRPKKLADMYGGGDD 829

RESULT 15
ADP26905
ID ADP26905 standard; protein; 829 AA.
XX AC ADP26905;
XX DT 26-AUG-2004 (first entry)
XX DE Human P-cadherin protein SEQ ID NO:6.
XX KW hair growth modulator; P-cadherin modulator; endocrine; depilatory;
XX KW gene therapy; antisense therapy; hair growth; alopecia; baldness;
XX KW unwanted hair growth; hirsutism;
XX KW hypotrichosis associated with juvenile macular dystrophy; HJMD; human;
XX KW P-cadherin.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Misc-difference 563
XX FT /note= "encoded by CAG"
XX PN EP1428893-A2.
XX PD 16-JUN-2004.
```

XX 10-OCT-2003; 2003EP-00256411.
 XX PF
 XX PR 15-OCT-2002; 2002US-0418163P.
 XX PR (SPRE/) SPRECHER E.
 XX PA (BERG/) BERGMAN R.
 XX PI Sprecher E, Bergman R;
 XX WPI; 2004-469945/45.
 DR N-PSDB; ADP26910.
 XX
 PT Identifying a hair growth modulator for treating alopecia and unwanted
 PT hair growth such as hirsutism, comprises identifying a P-cadherin
 PT modulator and testing whether the P-cadherin modulator is functional as a
 PT hair growth modulator.
 XX
 XX Disclosure; SEQ ID NO 6; 121bp; English.
 PS
 CC The present invention describes a method (M1) for identifying a hair
 CC growth modulator. (M1) comprises identifying a P-cadherin modulator, and
 CC testing whether the P-cadherin modulator is functional as a hair growth
 CC modulator. Also described: (I) a hair growth modulator (I) identified by
 CC (M1); and (2) a composition (II) for modulating hair growth, comprising,
 CC as an active ingredient, a P-cadherin modulator functional as a hair
 CC growth modulator. (I) and (II) have endocrine and depilatory activities,
 CC and can be used as hair growth modulators, P-cadherin function
 CC modulators, and in gene and antisense therapy. (M1) is useful for
 CC identifying a hair growth modulator. (I) is useful in a method of medical
 CC treatment. (I) or (II) is useful for modulating hair growth for non-
 CC therapeutic cosmetic purposes which involves administering to a subject,
 CC (I) or (II). (I) can be used in the manufacture of a medicament for the
 CC therapeutic modulation of hair growth. (I) or (II) is useful for treating
 CC alopecia (baldness) and unwanted hair growth such as hirsutism. (I) or
 CC (II) comprising P-cadherin inducer is useful for correction of hair loss
 CC in congenital hypotrichosis associated with juvenile macular dystrophy
 CC (HJMD) and other alopecia patients. The present sequence represents human
 CC P-cadherin, which is used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 829 AA;

Query Match 100.0%; Score 4369; DB 8; Length 829;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGPLASLLLLQVCMQCAASEPCRAVPREAEVTLGAGAEQEPQALGKVFMGCPG 60
 Db 1 MGLPRGPLASLLLLQVCMQCAASEPCRAVPREAEVTLGAGAEQEPQALGKVFMGCPG 60

Qy 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG 120
 Db 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDVWVAPISVPENG 120

Qy 121 KGPPFORLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREETAK 180
 Db 121 KGPPFORLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREETAK 180

Qy 181 YELFGHVAENGASVEDPMNISIIITVDQNDHKPKFTQDTFRGSVLEGVPGTSVMQVTAT 240
 Db 181 YELFGHVAENGASVEDPMNISIIITVDQNDHKPKFTQDTFRGSVLEGVPGTSVMQVTAT 240

Qy 241 DEDDAIYTYNGWAYSIHQEPKDPHDLMTIHRSTGTISVISGLDREKVPYTLTIQA 300
 Db 241 DEDDAIYTYNGWAYSIHQEPKDPHDLMTIHRSTGTISVISGLDREKVPYTLTIQA 300

Qy 301 TMDGDGSGTTTAVAVEILDANDNAPMFDPOKYEAHVPENAVGHEVQRLTVDLDAPNSP 360
 Db 301 TMDGDGSGTTTAVAVEILDANDNAPMFDPOKYEAHVPENAVGHEVQRLTVDLDAPNSP 360

Qy 361 AWRATYLMGGDGDHFTITTHPESNQGILTRKGLDFEAKNQHTLYVEVTNEAPFVLKL 420
 Db 361 AWRATYLMGGDGDHFTITTHPESNQGILTRKGLDFEAKNQHTLYVEVTNEAPFVLKL 420

Db 361 AWRATYLMGGDGDHFTITTHPESNQGILTRKGLDFEAKNQHTLYVEVTNEAPFVLKL 420
 Qy 421 PTSTATIVVHVEDVNEAPFVPPSKVVEVQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480
 Db 421 PTSTATIVVHVEDVNEAPFVPPSKVVEVQEGIPTGEPVCVYTAEDPDKENQKISYRILR 480
 Qy 481 DPAGWLAMPDSDQVTAAGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLLID 540
 Db 481 DPAGWLAMPDSDQVTAAGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLLID 540
 Qy 541 VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPHTSPPQAQLTDDSDIYWTAEVNEEGD 600
 Db 541 VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPHTSPPQAQLTDDSDIYWTAEVNEEGD 600
 Qy 601 TVVLSLKKFKLQDITYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWKGFIPLPVL 660
 Db 601 TVVLSLKKFKLQDITYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWKGFIPLPVL 660
 Qy 661 GAVLALLFLLLVLLLVLRKKIKIPEPLLLPEDDTRDNVFFYGGEGGEEDQYDITQIHR 720
 Db 661 GAVLALLFLLLVLLLVLRKKIKIPEPLLLPEDDTRDNVFFYGGEGGEEDQYDITQIHR 720
 Qy 721 GLEARPEVLRNDVAPTIIPTPMYRPRPANDPEIGNFIIENLKAANTDPTAPPYDTLLVF 780
 Db 721 GLEARPEVLRNDVAPTIIPTPMYRPRPANDPEIGNFIIENLKAANTDPTAPPYDTLLVF 780
 Qy 781 DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFPKKLADMYGGGDD 829
 Db 781 DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFPKKLADMYGGGDD 829

Search completed: December 12, 2005, 09:11:16
 Job time : 194 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2005, 09:08:03 ; Search time 167 Seconds
(without alignments)
2074.135 Million cell updates/sec

Title: US-10-696-639-39

Perfect score: 4369

Sequence: 1 MGLPRGPLASLLQLQVCWLQ.....NEWGRFKKLADMYGGEDD 829

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/prodata1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/prodata1/pubpaa/US10_PUBCOMB.pep.*
- 5: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/prodata1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4369	100.0	829	3	US-09-905-983-50
2	4369	100.0	829	3	US-09-916-849A-1
3	4369	100.0	829	3	US-09-975-723A-1
4	4369	100.0	829	4	US-10-174-677-21
5	4369	100.0	829	4	US-10-158-123-1
6	4369	100.0	829	4	US-10-295-027-783
7	4369	100.0	829	4	US-10-295-027-800
8	4369	100.0	829	4	US-10-295-027-863
9	4369	100.0	829	4	US-10-295-027-896
10	4369	100.0	829	4	US-10-058-270A-126
11	4369	100.0	829	5	US-10-643-795A-145
12	4369	100.0	829	5	US-10-578-160A-6
13	4369	100.0	829	5	US-10-723-860-2835
14	4369	100.0	829	5	US-10-696-639-39
15	4369	100.0	829	5	US-10-948-518-145
16	4369	100.0	829	5	US-10-756-149-5498
17	4369	100.0	829	6	US-11-037-713-22
18	4361	99.8	829	4	US-10-301-822-18
19	4361	99.8	829	4	US-10-295-027-338
20	4361	99.8	829	4	US-10-295-027-1246
21	4361	99.8	829	4	US-10-087-080-14
22	4361	99.8	829	4	US-10-229-345-16
23	4361	99.8	829	4	US-10-274-177-16
24	4361	99.8	829	4	US-10-650-112-16
25	4361	99.8	829	4	US-10-712-124-86
26	4361	99.8	829	6	US-11-041-788-16
27	2420	55.4	882	4	US-10-097-340-37

RESULT 1

US-09-905-983-50

; Sequence 50, Application US/09905983

; Patent No. US20020045591A1

; GENERAL INFORMATION:

; APPLICANT: Geiger, Benjamin

; APPLICANT: Ben-Ze'ev, Avri

; APPLICANT: Sadot, Eilat

; TITLE OF INVENTION: METHODS AND THERAPEUTIC COMPOSITIONS FOR TREATING CANCER

; FILE REFERENCE: 01/22326

; CURRENT APPLICATION NUMBER: US/09/905,983

; CURRENT FILING DATE: 2001-09-28

; NUMBER OF SEQ ID NOS: 60

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 50

; LENGTH: 829

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-905-983-50

Query Match 100.0%; Score 4369; DB 3; Length 829;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MGLPRGPLASLLQLQVCWLQCAASPCRAVFREAEVTLEAGGAEQEPGQALGKVFNGCPG	60
Db	1	MGLPRGPLASLLQLQVCWLQCAASPCRAVFREAEVTLEAGGAEQEPGQALGKVFNGCPG	60
QY	61	QEPALFSTDNDDFTVRNGETVQERSLKERNPLKIFPPSKRIILRRHKRDVWVAPISVPENG	120
Db	61	QEPALFSTDNDDFTVRNGETVQERSLKERNPLKIFPPSKRIILRRHKRDVWVAPISVPENG	120
QY	121	KGPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREIAK	180
Db	121	KGPFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREIAK	180
QY	181	YELFGHVASENSGASVEDPMNISIIITVDQNDHKPKFTQDTFRGSVLGVLPGTSMQVTTAT	240
Db	181	YELFGHVASENSGASVEDPMNISIIITVDQNDHKPKFTQDTFRGSVLGVLPGTSMQVTTAT	240
QY	241	DEDDAIYTYNGVAVSIHSQEPKDPHDLMTFTTHRSITGTSISVSSGLDRKVPYTTIOA	300
Db	241	DEDDAIYTYNGVAVSIHSQEPKDPHDLMTFTTHRSITGTSISVSSGLDRKVPYTTIOA	300
QY	301	TDMDGSGSTTTTAVAVVEILLDANDNAPMFPQKYEAHVPENAVGVHEVQRLTVDLDPNSP	360
Db	301	TDMDGSGSTTTTAVAVVEILLDANDNAPMFPQKYEAHVPENAVGVHEVQRLTVDLDPNSP	360
QY	361	AVRATYILMGDDGDHFTTTTHPESNQGLLTRKGLDFAKQOHTLYVEVTNEAPFVLKL	420
Db	361	AVRATYILMGDDGDHFTTTTHPESNQGLLTRKGLDFAKQOHTLYVEVTNEAPFVLKL	420

ALIGNMENTS

```
Db 361 AWRATYLMGGDDGDHFTITTHPESNOGILTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
Qy 421 PTSTATIVHVEDVNEAPFVPPSKVVEQEGIPTEGVCVYTAEDPDKENQKISYRILR 480
Db 421 PTSTATIVHVEDVNEAPFVPPSKVVEQEGIPTEGVCVYTAEDPDKENQKISYRILR 480
Qy 481 DPAGWLAMPDPSQGVTAAGTGLDREDEQFVRNNIYEVNVLAMNDGSPPTTGTGTLTLLID 540
Db 481 DPAGWLAMPDPSQGVTAAGTGLDREDEQFVRNNIYEVNVLAMNDGSPPTTGTGTLTLLID 540
Qy 541 VNDHGPVPEPRQITICNOSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPRQITICNOSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSSKKFLKQDHYDHLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Db 601 TVVLSSKKFLKQDHYDHLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Qy 661 GAVLALLFLLVLLVLLVLRKKRKEPILLPEDDTRDNVYFYGEGGEEQDQYDITQLHR 720
Db 661 GAVLALLFLLVLLVLLVLRKKRKEPILLPEDDTRDNVYFYGEGGEEQDQYDITQLHR 720
Qy 721 GLEARPEVLRNDVAPTIPTPMYRPRPANDPDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVLRNDVAPTIPTPMYRPRPANDPDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
Qy 781 DYEGSGSDAASLSLTSSASDQDQDYDYLNEGWSRPFKCLADMYGGGDD 829
Db 781 DYEGSGSDAASLSLTSSASDQDQDYDYLNEGWSRPFKCLADMYGGGDD 829

RESULT 2
US-09-916-849A-1
; Sequence 1, Application US/09916849A
; Publication No. US20030086934A1
; GENERAL INFORMATION:
; APPLICANT: Bostein, et al.
; TITLE OF INVENTION: Basal Markers in Breast Cancer and Related Reagents
; FILE REFERENCE: 2002850-0024
; CURRENT APPLICATION NUMBER: US/09/916-849A
; PRIOR FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Sequence of
; OTHER INFORMATION: Cadherin 3
US-09-916-849A-1

Query Match 100.0%; Score 4369; DB 3; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGPLASLLILQVCWLQCAASEPCRAVFREAEVTLGAGAEQEPQALGKVFMCPCG 60
Db 1 MGLPRGPLASLLILQVCWLQCAASEPCRAVFREAEVTLGAGAEQEPQALGKVFMCPCG 60
Qy 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHRKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHRKRDWVAPISVPENG 120
Qy 121 KGPFPQRLNQLKSNKDRDTKIFYITGPGADSPPEGFAVEKETGWLILNKLPLDREBIAK 180
Db 121 KGPFPQRLNQLKSNKDRDTKIFYITGPGADSPPEGFAVEKETGWLILNKLPLDREBIAK 180
Qy 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSLVGLVPGTSVMQVTAT 240
Db 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSLVGLVPGTSVMQVTAT 240
```

```
Qy 241 DEDDAIYTYNGVWAYGIHSQEPKDPHDLMTIHRSTGTISVISSSGLDREKVPYTLTIOA 300
Db 241 DEDDAIYTYNGVWAYGIHSQEPKDPHDLMTIHRSTGTISVISSSGLDREKVPYTLTIOA 300
Qy 301 TMDGSGSTTTAVAVVEILDANDNAPMFPDQPKYEAHVPENAVGHVQRLTVTDLDPNSP 360
Db 301 TMDGSGSTTTAVAVVEILDANDNAPMFPDQPKYEAHVPENAVGHVQRLTVTDLDPNSP 360
Qy 361 AWRATYLMGGDDGDHFTITTHPESNOGILTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
Db 361 AWRATYLMGGDDGDHFTITTHPESNOGILTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
Qy 421 PTSTATIVHVEDVNEAPFVPPSKVVEQEGIPTEGVCVYTAEDPDKENQKISYRILR 480
Db 421 PTSTATIVHVEDVNEAPFVPPSKVVEQEGIPTEGVCVYTAEDPDKENQKISYRILR 480
Qy 481 DPAGWLAMPDPSQGVTAAGTGLDREDEQFVRNNIYEVNVLAMNDGSPPTTGTGTLTLLID 540
Db 481 DPAGWLAMPDPSQGVTAAGTGLDREDEQFVRNNIYEVNVLAMNDGSPPTTGTGTLTLLID 540
Qy 541 VNDHGPVPEPRQITICNOSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPRQITICNOSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSSKKFLKQDHYDHLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Db 601 TVVLSSKKFLKQDHYDHLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILLPVL 660
Qy 661 GAVLALLFLLVLLVLLVLRKKRKEPILLPEDDTRDNVYFYGEGGEEQDQYDITQLHR 720
Db 661 GAVLALLFLLVLLVLLVLRKKRKEPILLPEDDTRDNVYFYGEGGEEQDQYDITQLHR 720
Qy 721 GLEARPEVLRNDVAPTIPTPMYRPRPANDPDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVLRNDVAPTIPTPMYRPRPANDPDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
Qy 781 DYEGSGSDAASLSLTSSASDQDQDYDYLNEGWSRPFKCLADMYGGGDD 829
Db 781 DYEGSGSDAASLSLTSSASDQDQDYDYLNEGWSRPFKCLADMYGGGDD 829

RESULT 3
US-09-975-723A-1
; Sequence 1, Application US/09975723A
; Publication No. US20030108529A1
; GENERAL INFORMATION:
; APPLICANT: Nackman, Gary
; TITLE OF INVENTION: Improvement of Endothelial Cell-Cell
; FILE REFERENCE: 601-1-101N
; CURRENT APPLICATION NUMBER: US/09/975,723A
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/241,216
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/243,693
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-975-723A-1

Query Match 100.0%; Score 4369; DB 3; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGPLASLLILQVCWLQCAASEPCRAVFREAEVTLGAGAEQEPQALGKVFMCPCG 60
Db 1 MGLPRGPLASLLILQVCWLQCAASEPCRAVFREAEVTLGAGAEQEPQALGKVFMCPCG 60
```


Qy	61	QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDWVAPISVPENG	120
Db	61	QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDWVAPISVPENG	120
Qy	121	KGPPORLNOLKSNKORDTKIFYSITGPGADSPPEGVAVEKETGWLKLNKPLDREEIAK	180
Db	121	KGPPORLNOLKSNKORDTKIFYSITGPGADSPPEGVAVEKETGWLKLNKPLDREEIAK	180
Qy	181	YELFGHVAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSLVGLVPGTSMQVTTAT	240
Db	181	YELFGHVAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSLVGLVPGTSMQVTTAT	240
Qy	241	DEDDAIYTYNGVAVSIHSQEPKDPHDLMTIHRSTGTISVISGLDREKVPYTYLTIQA	300
Db	241	DEDDAIYTYNGVAVSIHSQEPKDPHDLMTIHRSTGTISVISGLDREKVPYTYLTIQA	300
Qy	301	TDMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVQRLTVDLDAPNSP	360
Db	301	TDMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVQRLTVDLDAPNSP	360
Qy	361	AWRATYLMGGDGDHFTIITHPESNQILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420
Db	361	AWRATYLMGGDGDHFTIITHPESNQILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420
Qy	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRIILR	480
Db	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRIILR	480
Qy	481	DPAGWLAMPDPSQVTAAGTGLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLTLID	540
Db	481	DPAGWLAMPDPSQVTAAGTGLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLTLID	540
Qy	541	VNDHGPVPEPRQITTCNQSPVRHLNITDKDLSPHTSPPFOAQLTDDSDIYWTAEVNEEGD	600
Db	541	VNDHGPVPEPRQITTCNQSPVRHLNITDKDLSPHTSPPFOAQLTDDSDIYWTAEVNEEGD	600
Qy	601	TVVLSLKKFLKQDQTYDVHLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILPVL	660
Db	601	TVVLSLKKFLKQDQTYDVHLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILPVL	660
Qy	661	GAVLALLFLLVLLVLLVLRKKRKIKPEPLLPEDDTRDNVFFYEGEGGEDDQDYITQLHR	720
Db	661	GAVLALLFLLVLLVLLVLRKKRKIKPEPLLPEDDTRDNVFFYEGEGGEDDQDYITQLHR	720
Qy	721	GLEARPEVLRNDVAPTIIPTMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF	780
Db	721	GLEARPEVLRNDVAPTIIPTMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF	780
Qy	781	DYEGSGDAASLSLTSASDQDQDYDYLNEWGSRFKKLADMYGGGEDD	829
Db	781	DYEGSGDAASLSLTSASDQDQDYDYLNEWGSRFKKLADMYGGGEDD	829

RESULT 4
US-10-174-677-21
; Sequence 21, Application US/10174677
; Publication No. US20030190704A1
; GENERAL INFORMATION:
; APPLICANT: Xie, Ting
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ANCHORING STEM CELLS IN A MICROENVIR
; FILE REFERENCES: 40716 (IP-012)
; CURRENT APPLICATION NUMBER: US/10/174,677
; CURRENT FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-677-21

Query Match 100.0%; Score 4369; DB 4; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;

Matches	829;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;						
Qy	1	MGLPRGPLASLLLLQVCMWLQCAASEPCRAVFR	EA	EV	TL	EA	GA	EP	GA	LG	KV	FM	GC	PG	60
Db	1	MGLPRGPLASLLLLQVCMWLQCAASEPCRAVFR	EA	EV	TL	EA	GA	EP	GA	LG	KV	FM	GC	PG	60
Qy	61	QBPALFSTDNDPFTVRNGETVQERRSI	KERNPLKIPPSKRILRRHKRDWVAPISVPENG	120											
Db	61	QBPALFSTDNDPFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDWVAPISVPENG	120												
Qy	121	KGPPFORLNOLKSNKORDTKIFYSITGPGADSP	PEGVFAVEKETGWLKLNKPLDREEIAK	180											
Db	121	KGPPFORLNOLKSNKORDTKIFYSITGPGADSP	PEGVFAVEKETGWLKLNKPLDREEIAK	180											
Qy	181	YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDT	FRGSLVGLVPGTSMQVTTAT	240											
Db	181	YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDT	FRGSLVGLVPGTSMQVTTAT	240											
Qy	241	DEDDAIYTYNGVAVSIHSQEPKDPHDLMTIHRSTGTISVIS	SGLDREKVPYTYLTIQA	300											
Db	241	DEDDAIYTYNGVAVSIHSQEPKDPHDLMTIHRSTGTISVIS	SGLDREKVPYTYLTIQA	300											
Qy	301	TDMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAHV	PENAVGHEVQRLTVDLDAPNSP	360											
Db	301	TDMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAHV	PENAVGHEVQRLTVDLDAPNSP	360											
Qy	361	AWRATYLMGGDDGHDFTIITHPESNOGILTTTRKGLD	FEAKNQHTLYVEVTNEAPFVLKL	420											
Db	361	AWRATYLMGGDDGHDFTIITHPESNOGILTTTRKGLD	FEAKNQHTLYVEVTNEAPFVLKL	420											
Qy	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEP	VCVYTAEDPDKENQKISYRIILR	480											
Db	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEP	VCVYTAEDPDKENQKISYRIILR	480											
Qy	481	DPAGWLAMPDPSQVTAAGTGLDREDEQFVRNNIYEV	WVLAMDNQSPPTTGTGTLTLTLID	540											
Db	481	DPAGWLAMPDPSQVTAAGTGLDREDEQFVRNNIYEV	WVLAMDNQSPPTTGTGTLTLTLID	540											
Qy	541	VNDHGPVPEPRQITTCNQSPVRHLNITDKDLSPHTS	PPFOAQLTDDSDIYMTAEVNEEGD	600											
Db	541	VNDHGPVPEPRQITTCNQSPVRHLNITDKDLSPHTS	PPFOAQLTDDSDIYMTAEVNEEGD	600											
Qy	601	TVVLSLKKFLKQDQTYDVHLSDHGNKEQLTVIRAT	VCDCHGHVETCPGPKWGGFILPVL	660											
Db	601	TVVLSLKKFLKQDQTYDVHLSDHGNKEQLTVIRAT	VCDCHGHVETCPGPKWGGFILPVL	660											
Qy	661	GAVLALLFLLVLLVLLVLRKKRKIKPEPLLPEDD	TRDNVFFYEGEGGEDDQDYITQLHR	720											
Db	661	GAVLALLFLLVLLVLLVLRKKRKIKPEPLLPEDD	TRDNVFFYEGEGGEDDQDYITQLHR	720											
Qy	721	GLEARPEVLRNDVAPTIIPTMYRPRPANDPDEIGN	FIENLKAANTDPTAPPYDTLLVF	780											
Db	721	GLEARPEVLRNDVAPTIIPTMYRPRPANDPDEIGN	FIENLKAANTDPTAPPYDTLLVF	780											
Qy	781	DYEGSGDAASLSLTSASDQDQDYDYLNEWGSR	FKKLADMYGGGEDD	829											
Db	781	DYEGSGDAASLSLTSASDQDQDYDYLNEWGSR	FKKLADMYGGGEDD	829											

RESULT 5

US-10-158-123-1

Sequence 1, Application US/10158123

Publication No. US20030194406A1

GENERAL INFORMATION:

APPLICANT: Reinhard, Christoph

APPLICANT: Klinger, Julie

APPLICANT: Jefferson, Ann

APPLICANT: Escobedo, Jaime

APPLICANT: Randazzo, Fillipo

APPLICANT: Winter, Jill

APPLICANT: Goodson, Robert

APPLICANT: Qi, Weimin

TITLE OF INVENTION: P-Cadherin as a Target for Anti-Cancer

RESULT 5
US-10-158-123-1
; Sequence 1, Application US/10158123
; Publication No. US20030194406A1
; GENERAL INFORMATION:
; APPLICANT: Reinhard, Christoph
; APPLICANT: Klinger, Julie
; APPLICANT: Jefferson, Ann
; APPLICANT: Escobedo, Jaime
; APPLICANT: Randazzo, Fillipo
; APPLICANT: Winter, Jill
; APPLICANT: Goodson, Robert
; APPLICANT: Qi, Weimin
; TITLE OF INVENTION: P-Cadherin as a Target for Anti-Cancer

```
; TITLE OF INVENTION: Therapy
; FILE REFERENCE: 35784/258994
; CURRENT APPLICATION NUMBER: US/10/158,123
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/294,225
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-158-123-1

Query Match      100.0%; Score 4369; DB 4; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLEAGGAEQEQALGKVFMGCPG 60
Db 1 MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLEAGGAEQEQALGKVFMGCPG 60

Qy 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120

Qy 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120

Qy 121 KGPFPQRLNQLKSKNDRTKIFYSTITGPGADSPPEGVFAVEKETGWLKLLNKPLDREIAK 180
Db 121 KGPFPQRLNQLKSKNDRTKIFYSTITGPGADSPPEGVFAVEKETGWLKLLNKPLDREIAK 180

Qy 121 KGPFPQRLNQLKSKNDRTKIFYSTITGPGADSPPEGVFAVEKETGWLKLLNKPLDREIAK 180
Db 121 KGPFPQRLNQLKSKNDRTKIFYSTITGPGADSPPEGVFAVEKETGWLKLLNKPLDREIAK 180

Qy 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVTAT 240
Db 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVTAT 240

Qy 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVTAT 240
Db 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVTAT 240

Qy 241 DEDDAIYTYNGVWAYSIHSGQPKDPLDMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
Db 241 DEDDAIYTYNGVWAYSIHSGQPKDPLDMFTIHRSTGTISVISGLDREKVPYTLTIOA 300

; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 783
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-295-027-783

Query Match      100.0%; Score 4369; DB 4; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLEAGGAEQEQALGKVFMGCPG 60
Db 1 MGLPRGPLASLLLLQVCWLQCAASEPCRAVFREAEVTLEAGGAEQEQALGKVFMGCPG 60

Qy 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120

Qy 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHKRDWVAPISVPENG 120

Qy 121 KGPFPQRLNQLKSKNDRTKIFYSTITGPGADSPPEGVFAVEKETGWLKLLNKPLDREIAK 180
Db 121 KGPFPQRLNQLKSKNDRTKIFYSTITGPGADSPPEGVFAVEKETGWLKLLNKPLDREIAK 180

Qy 121 KGPFPQRLNQLKSKNDRTKIFYSTITGPGADSPPEGVFAVEKETGWLKLLNKPLDREIAK 180
Db 121 KGPFPQRLNQLKSKNDRTKIFYSTITGPGADSPPEGVFAVEKETGWLKLLNKPLDREIAK 180

Qy 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVTAT 240
Db 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVTAT 240

Qy 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVTAT 240
Db 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVTAT 240

Qy 241 DEDDAIYTYNGVWAYSIHSGQPKDPLDMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
Db 241 DEDDAIYTYNGVWAYSIHSGQPKDPLDMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
```

Db 241 DEDDAIYVNGVAYSHSQEPKDPHDLMTTHRSCTGTISVSSGLDREKVPYTLTIOA 300
Qy 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAAHVPENAVGHEVQRLTVDLDPNSP 360
Db 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAAHVPENAVGHEVQRLTVDLDPNSP 360
Qy 361 AWRATYLMGDDGDHFTTTHPESNOGILTRKGLDPEAKNOHTLYVEVTNEAPFVLK 420
Db 361 AWRATYLMGDDGDHFTTTHPESNOGILTRKGLDPEAKNOHTLYVEVTNEAPFVLK 420
Qy 421 PTSTATIVHVEDVNEAPVFPVPSKVVEOEGIPTEGPEVCVYTAEDPDKENOKISYRI 480
Db 421 PTSTATIVHVEDVNEAPVFPVPSKVVEOEGIPTEGPEVCVYTAEDPDKENOKISYRI 480
Qy 481 DPAGWLAMPDSDGQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGLTLLTID 540
Db 481 DPAGWLAMPDSDGQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGLTLLTID 540
Qy 541 VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSLKKFELKQDVTYDVHLSLSDHGNKEQLTVIRATVCDCHGVETCPGPKGGFLLPVL 660
Db 601 TVVLSLKKFELKQDVTYDVHLSLSDHGNKEQLTVIRATVCDCHGVETCPGPKGGFLLPVL 660
Qy 661 GAVLALLFLLVLLVLLVRKRRKIKEPLLPEDDTRDNVFFYEGEGEEDQDYDTQLHR 720
Db 661 GAVLALLFLLVLLVLLVRKRRKIKEPLLPEDDTRDNVFFYEGEGEEDQDYDTQLHR 720
Qy 721 GLEARPEVLRNDVAPTIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVLRNDVAPTIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
Qy 781 DYEAGSDDAASLSLTSASDQDYDYLNEGSRFKKLADMYGGGEDD 829
Db 781 DYEAGSDDAASLSLTSASDQDYDYLNEGSRFKKLADMYGGGEDD 829

RESULT 7

US-10-295-027-800
; Sequence 800, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08

; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 800
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-800

Query Match 100.0%; Score 4369; DB 4; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLPRGPLASLLLLQVCMLQCAASBPCRAVFPREAEVTLBAGGAEOPGQALGKVFNGCPG 60
Db 1 MGLPRGPLASLLLLQVCMLQCAASBPCRAVFPREAEVTLBAGGAEOPGQALGKVFNGCPG 60
Qy 61 QEPALFSTDNDDFTVRNGETVQERSLKERNPLKIPPSKRIILRRHKRDVAVAPISVPENG 120
Db 61 QEPALFSTDNDDFTVRNGETVQERSLKERNPLKIPPSKRIILRRHKRDVAVAPISVPENG 120
Qy 121 KGPFQRLNQLKSNKDRDRTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREEIAK 180
Db 121 KGPFQRLNQLKSNKDRDRTKIFYSITGPGADSPPEGVFAVEKETGWLKLNKPLDREEIAK 180
Qy 181 YELFHGAVSENGASVEDPMNISIIITDQNDHKKPTQDTFRGSLVLEGLVPGTSMQVAT 240
Db 181 YELFHGAVSENGASVEDPMNISIIITDQNDHKKPTQDTFRGSLVLEGLVPGTSMQVAT 240
Qy 241 DEDDAIYVNGVAYSHSQEPKDPHDLMTTHRSCTGTISVSSGLDREKVPYTLTIOA 300
Db 241 DEDDAIYVNGVAYSHSQEPKDPHDLMTTHRSCTGTISVSSGLDREKVPYTLTIOA 300
Qy 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAAHVPENAVGHEVQRLTVDLDPNSP 360
Db 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOKYEAAHVPENAVGHEVQRLTVDLDPNSP 360
Qy 361 AWRATYLMGDDGDHFTTTHPESNOGILTRKGLDPEAKNOHTLYVEVTNEAPFVLK 420
Db 361 AWRATYLMGDDGDHFTTTHPESNOGILTRKGLDPEAKNOHTLYVEVTNEAPFVLK 420
Qy 421 PTSTATIVHVEDVNEAPVFPVPSKVVEOEGIPTEGPEVCVYTAEDPDKENOKISYRI 480
Db 421 PTSTATIVHVEDVNEAPVFPVPSKVVEOEGIPTEGPEVCVYTAEDPDKENOKISYRI 480
Qy 481 DPAGWLAMPDSDGQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGLTLLTID 540
Db 481 DPAGWLAMPDSDGQVTAAGTLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGLTLLTID 540
Qy 541 VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSLKKFELKQDVTYDVHLSLSDHGNKEQLTVIRATVCDCHGVETCPGPKGGFLLPVL 660
Db 601 TVVLSLKKFELKQDVTYDVHLSLSDHGNKEQLTVIRATVCDCHGVETCPGPKGGFLLPVL 660
Qy 661 GAVLALLFLLVLLVLLVRKRRKIKEPLLPEDDTRDNVFFYEGEGEEDQDYDTQLHR 720
Db 661 GAVLALLFLLVLLVLLVRKRRKIKEPLLPEDDTRDNVFFYEGEGEEDQDYDTQLHR 720
Qy 721 GLEARPEVLRNDVAPTIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVLRNDVAPTIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
Qy 781 DYEAGSDDAASLSLTSASDQDYDYLNEGSRFKKLADMYGGGEDD 829
Db 781 DYEAGSDDAASLSLTSASDQDYDYLNEGSRFKKLADMYGGGEDD 829


```

; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 896
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-295-027-896

Query Match      100.0%; Score 4369; DB 4; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1  MGLPRGPLASILLQVCWLQCAASEPCRAVFREAETLEAGGAEQEQALGKGVFMC 60
Db  1  MGLPRGPLASILLQVCWLQCAASEPCRAVFREAETLEAGGAEQEQALGKGVFMC 60

Qy  61  QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPSKRILRRHKRDVWVAPISV 120
Db  61  QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPSKRILRRHKRDVWVAPISV 120

Qy  121  KGPPPPQRLNQLKSNKDRDTKIFYISITGPGADSPPEGVFAVEKETGWLNLNKPLOREE 180
Db  121  KGPPPPQRLNQLKSNKDRDTKIFYISITGPGADSPPEGVFAVEKETGWLNLNKPLOREE 180

Qy  181  YELFGHAVSENGASVEDPMNISIIVTQNDHKPKFTQDTRFRGSVLEGLPGTSMVQV 240
Db  181  YELFGHAVSENGASVEDPMNISIIVTQNDHKPKFTQDTRFRGSVLEGLPGTSMVQV 240

Qy  241  DEDDAIYTYNGVVAYSIHSQBPDPDLMTFIHRSTGTISVISSSLGDRKVPYETL 300
Db  241  DEDDAIYTYNGVVAYSIHSQBPDPDLMTFIHRSTGTISVISSSLGDRKVPYETL 300

Qy  301  TDMDGDSGTTTAVAVELDANDNAPMFDPOKYEAHVPENAVGHEVORLTVTDLDA 360
Db  301  TDMDGDSGTTTAVAVELDANDNAPMFDPOKYEAHVPENAVGHEVORLTVTDLDA 360

Qy  361  AWRATYILMGDGDGHFTITTHPSNQGLTTRKGLDPEAKNOHTLVVEVNEAPFVL 420
Db  361  AWRATYILMGDGDGHFTITTHPSNQGLTTRKGLDPEAKNOHTLVVEVNEAPFVL 420

Qy  421  PTSTATIVHVHVEDVNEAPFVPPSKVVEVQSGIPTGEPVCVYTAEDPDKENQIS 480
Db  421  PTSTATIVHVHVEDVNEAPFVPPSKVVEVQSGIPTGEPVCVYTAEDPDKENQIS 480

Qy  481  DPAGWLAMPDSGGVTAAGTLDRDEDEQFVRNNIYEVWVLANDNGSPPTGTGTL 540
Db  481  DPAGWLAMPDSGGVTAAGTLDRDEDEQFVRNNIYEVWVLANDNGSPPTGTGTL 540

Qy  541  VNDHGVPPEPQIITCNQSPVRHVNLITDKLSPHTSPFOAQLTDDSDIYWTAEV 600
Db  541  VNDHGVPPEPQIITCNQSPVRHVNLITDKLSPHTSPFOAQLTDDSDIYWTAEV 600

Qy  601  TVVLSLKKFLKQDITYVHLSLSDSHGKNEQLTVIRATVCDCHGHEDEDDYDI 660
Db  601  TVVLSLKKFLKQDITYVHLSLSDSHGKNEQLTVIRATVCDCHGHEDEDDYDI 660

Qy  661  GAVLALLFLLLVLLVLLVRKKGIKEPILLPDDTRDNVFFYGGEGGEGDEDDY 720
Db  661  GAVLALLFLLLVLLVLLVRKKGIKEPILLPDDTRDNVFFYGGEGGEGDEDDY 720

Qy  721  GLEARPEVLKNDVAPITPTPMYRPPRANPEIGNFIIENLKAANTDPTAPPYD 780
Db  721  GLEARPEVLKNDVAPITPTPMYRPPRANPEIGNFIIENLKAANTDPTAPPYD 780

Qy  781  DYEGSGDAASLSLTSASDQDDYDYNLNEWGSRFKKLADMYGGGDD 829

```

Qy	481	DPAGWLAMPDPSGOVTAVGTLDRDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLLID	540
Db	481	DPAGWLAMPDPSGOVTAVGTLDRDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLLID	540
Qy	541	VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD	600
Db	541	VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD	600
Qy	601	TVVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKGSGFILPVL	660
Db	601	TVVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKGSGFILPVL	660
Qy	661	GAVALALLFLLLVLLVLRKKRKEPILLPDDTRDNVFFYGGEGGEEQDDYDITOLHR	720
Db	661	GAVALALLFLLLVLLVLRKKRKEPILLPDDTRDNVFFYGGEGGEEQDDYDITOLHR	720
Qy	721	GLEARPEVLRNDVAPTIIPTPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF	780
Db	721	GLEARPEVLRNDVAPTIIPTPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF	780
Qy	781	DYEGSGSDAASLSLTSASDQDDYDYLNEWGSRFKKLADMYGGGDD	829
Db	781	DYEGSGSDAASLSLTSASDQDDYDYLNEWGSRFKKLADMYGGGDD	829
RESULT 11			
US-10-643-795A-145			
; Sequence 145, Application US/10643795A			
; Publication No. US20040241703A1			
; GENERAL INFORMATION:			
; APPLICANT: FREDERIC J. DESAUVAGE			
; APPLICANT: GRETCHEN FRANTZ			
; APPLICANT: KENNETH J. HILLAN			
; APPLICANT: PAUL POLAKIS			
; APPLICANT: ANDREW POLSON			
; APPLICANT: VICTORIA SMITH			
; APPLICANT: SUSAN D. SPENCER			
; APPLICANT: THOMAS D. WU			
; APPLICANT: ZEMIN ZHANG			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND			
; TITLE OF INVENTION: TREATMENT OF TUMOR			
; FILE REFERENCE: P5026R1-US			
; CURRENT APPLICATION NUMBER: US/10/643,795A			
; CURRENT FILING DATE: 2003-08-19			
; PRIOR FILING DATE: 2002-08-19			
; PRIOR APPLICATION NUMBER: US 60/404,809			
; PRIOR FILING DATE: 2002-08-19			
; PRIOR APPLICATION NUMBER: US 60/405,645			
; PRIOR FILING DATE: 2002-08-21			
; PRIOR APPLICATION NUMBER: US 60/413,192			
; PRIOR FILING DATE: 2002-09-23			
; PRIOR APPLICATION NUMBER: US 60/419,008			
; PRIOR FILING DATE: 2002-10-15			
; PRIOR APPLICATION NUMBER: US 60/426,847			
; PRIOR FILING DATE: 2002-11-15			
; PRIOR APPLICATION NUMBER: US 60/484,959			
; PRIOR FILING DATE: 2003-07-02			
; NUMBER OF SEQ ID NOS: 158			
; SEQ ID NO 145			
; LENGTH: 829			
; TYPE: PRT			
; ORGANISM: Homo sapien			
US-10-643-795A-145			
Query Match 100.0%; Score 4369; DB 5; Length 829;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MGLPRGPLASLLLLQVCHLOCAASPCPRAVREAVTLEAGAGSEPCQALGKVMGCPG	60
Db	1	MGLPRGPLASLLLLQVCHLOCAASPCPRAVREAVTLEAGAGSEPCQALGKVMGCPG	60
Qy	61	QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHRKRDWVAPISVPENG	120

Db	61	QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRILRRHRKRDWVAPISVPENG	120
Qy	121	KGFPORLNQKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDRREIAK	180
Db	121	KGFPORLNQKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGMLLNKPLDRREIAK	180
Qy	181	YELFGHAVSENGASVEDPMNISIIVTDQNDHDKPKFTQDTFRGSVLGVLPGTSMQVTAT	240
Db	181	YELFGHAVSENGASVEDPMNISIIVTDQNDHDKPKFTQDTFRGSVLGVLPGTSMQVTAT	240
Qy	241	DEDDAIYTYNGVAVSIHQEPKDPHLMFTIHRSTGTISVISGSLDREKVPYTLTIOA	300
Db	241	DEDDAIYTYNGVAVSIHQEPKDPHLMFTIHRSTGTISVISGSLDREKVPYTLTIOA	300
Qy	301	TDMDGSGSTTTAVAVVEILDANDANPMDPOKQEAHVPENAVGHVQRLTVTOLDADNSP	360
Db	301	TDMDGSGSTTTAVAVVEILDANDANPMDPOKQEAHVPENAVGHVQRLTVTOLDADNSP	360
Qy	361	AWRATYLLIMGDDGDHFTITTHPESNQGLITTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420
Db	361	AWRATYLLIMGDDGDHFTITTHPESNQGLITTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420
Qy	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEVOEGIPTGEPVCVYTAEDPDKENQKISYRILR	480
Db	421	PTSTATIVVHVEDVNEAPVFPVPSKVVEVOEGIPTGEPVCVYTAEDPDKENQKISYRILR	480
Qy	481	DPAGWLAMPDPSGOVTAVGTLDRDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLLID	540
Db	481	DPAGWLAMPDPSGOVTAVGTLDRDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLLID	540
Qy	541	VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD	600
Db	541	VNDHGPVPEPRQITICNQSPVRHLNITDKDLSPTSPFQAQLTDDSDIYWTAEVNEEGD	600
Qy	601	TVVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKGSGFILPVL	660
Db	601	TVVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKGSGFILPVL	660
Qy	661	GAVALALLFLLLVLLVLRKKRKEPILLPDDTRDNVFFYGGEGGEEQDDYDITOLHR	720
Db	661	GAVALALLFLLLVLLVLRKKRKEPILLPDDTRDNVFFYGGEGGEEQDDYDITOLHR	720
Qy	721	GLEARPEVLRNDVAPTIIPTPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF	780
Db	721	GLEARPEVLRNDVAPTIIPTPMYRPRPANDPDEIGNFIENLKAANTDPTAPPYDTLLVF	780
Qy	781	DYEGSGSDAASLSLTSASDQDDYDYLNEWGSRFKKLADMYGGGDD	829
Db	781	DYEGSGSDAASLSLTSASDQDDYDYLNEWGSRFKKLADMYGGGDD	829
RESULT 12			
US-10-678-160A-6			
; Sequence 6, Application US/10678160A			
; Publication No. US20040247555A1			
; GENERAL INFORMATION:			
; APPLICANT: Sprecher, Eli			
; APPLICANT: Bergman, Reuven			
; TITLE OF INVENTION: METHODS OF AND COMPOSITIONS FOR MODULATING HAIR GROWTH VIA			
; TITLE OF INVENTION: P-CADHERIN MODULATORS			
; FILE REFERENCE: 26465			
; CURRENT APPLICATION NUMBER: US/10/678,160A			
; CURRENT FILING DATE: 2003-10-06			
; NUMBER OF SEQ ID NOS: 75			
; SOFTWARE: Patent in version 3.1			
; SEQ ID NO 6			
; LENGTH: 829			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-678-160A-6			
Query Match 100.0%; Score 4369; DB 5; Length 829;			


```
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLPRGLASLLLLQVLCWLCQAASEPCRAVFREAEVTLLEAGGAEQEPGQALGKVFVCGPG 60
    |||||
Db 1 MGLPRGLASLLLLQVLCWLCQAASEPCRAVFREAEVTLLEAGGAEQEPGQALGKVFVCGPG 60
    |||||

QY 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRIILRRHKRDVWVAPISVPENG 120
    |||||
Db 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRIILRRHKRDVWVAPISVPENG 120
    |||||

QY 121 KGPFFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKKPLDREIEIAK 180
    |||||
Db 121 KGPFFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKKPLDREIEIAK 180
    |||||

QY 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTPRGSVLEGLPGTSVMQVTAT 240
    |||||
Db 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTPRGSVLEGLPGTSVMQVTAT 240
    |||||

QY 241 DEDDAIYTVNGVAVYSIHSQEPKDPHDLMTIHRSTGTISVISGLDRKVPYETLTIOA 300
    |||||
Db 241 DEDDAIYTVNGVAVYSIHSQEPKDPHDLMTIHRSTGTISVISGLDRKVPYETLTIOA 300
    |||||

QY 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOQYEAHVPENAVGVHEVQRLTVTDLDPNSP 360
    |||||
Db 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOQYEAHVPENAVGVHEVQRLTVTDLDPNSP 360
    |||||

QY 361 AWRATYILMGDDGDHFTITTHPESNQILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
    |||||
Db 361 AWRATYILMGDDGDHFTITTHPESNQILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
    |||||

QY 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRIILR 480
    |||||
Db 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRIILR 480
    |||||

QY 481 DPAGWLAMDPSQVTAAGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLTLD 540
    |||||
Db 481 DPAGWLAMDPSQVTAAGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLTLD 540
    |||||

QY 541 VNDHGPVPEPRQITICNQSPVRHVLNITDKDLSPTSPPQAQLTDDSDIYWTAEVNEEGD 600
    |||||
Db 541 VNDHGPVPEPRQITICNQSPVRHVLNITDKDLSPTSPPQAQLTDDSDIYWTAEVNEEGD 600
    |||||

QY 601 TVVLSLKKFLKQDPTDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGWKGGFLLPVL 660
    |||||
Db 601 TVVLSLKKFLKQDPTDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGWKGGFLLPVL 660
    |||||

QY 661 GAVLALLFLLVLLVLLVLRKKRKIKPEPLLPEDDTRDNVFFYEGEGGEDDQDYDTQLHR 720
    |||||
Db 661 GAVLALLFLLVLLVLLVLRKKRKIKPEPLLPEDDTRDNVFFYEGEGGEDDQDYDTQLHR 720
    |||||

QY 721 GLEARPEVLRNDVAPTIIPTPMYRPRANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
    |||||
Db 721 GLEARPEVLRNDVAPTIIPTPMYRPRANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
    |||||

QY 781 DYEGSGSDAASLSLSTSSASDQDQDYDYLNEWGSFRFKKLADMYGGEDD 829
    |||||
Db 781 DYEGSGSDAASLSLSTSSASDQDQDYDYLNEWGSFRFKKLADMYGGEDD 829
    |||||
```

RESULT 13

```
US-10-723-860-2835
; Sequence 2835, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NFUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
```

```
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2835
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-2835

Query Match 100.0%; Score 4369; DB 5; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLPRGLASLLLLQVLCWLCQAASEPCRAVFREAEVTLLEAGGAEQEPGQALGKVFVCGPG 60
    |||||
Db 1 MGLPRGLASLLLLQVLCWLCQAASEPCRAVFREAEVTLLEAGGAEQEPGQALGKVFVCGPG 60
    |||||

QY 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRIILRRHKRDVWVAPISVPENG 120
    |||||
Db 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRIILRRHKRDVWVAPISVPENG 120
    |||||

QY 121 KGPFFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKKPLDREIEIAK 180
    |||||
Db 121 KGPFFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKKPLDREIEIAK 180
    |||||

QY 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTPRGSVLEGLPGTSVMQVTAT 240
    |||||
Db 181 YELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKFTQDTPRGSVLEGLPGTSVMQVTAT 240
    |||||

QY 241 DEDDAIYTVNGVAVYSIHSQEPKDPHDLMTIHRSTGTISVISGLDRKVPYETLTIOA 300
    |||||
Db 241 DEDDAIYTVNGVAVYSIHSQEPKDPHDLMTIHRSTGTISVISGLDRKVPYETLTIOA 300
    |||||

QY 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOQYEAHVPENAVGVHEVQRLTVTDLDPNSP 360
    |||||
Db 301 TMDGSGSTTTAVAVVEILDANDNAPMFDPOQYEAHVPENAVGVHEVQRLTVTDLDPNSP 360
    |||||

QY 361 AWRATYILMGDDGDHFTITTHPESNQILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
    |||||
Db 361 AWRATYILMGDDGDHFTITTHPESNQILTTTRKGLDFAKNQHTLYVEVTNEAPFVLKL 420
    |||||

QY 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRIILR 480
    |||||
Db 421 PTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTGEPVCVYTAEDPDKENQKISYRIILR 480
    |||||

QY 481 DPAGWLAMDPSQVTAAGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLTLD 540
    |||||
Db 481 DPAGWLAMDPSQVTAAGTLDREDEQFVRNNIYEVWVLAMDNQSPPTTGTGTLTLTLD 540
    |||||

QY 541 VNDHGPVPEPRQITICNQSPVRHVLNITDKDLSPTSPPQAQLTDDSDIYWTAEVNEEGD 600
    |||||
Db 541 VNDHGPVPEPRQITICNQSPVRHVLNITDKDLSPTSPPQAQLTDDSDIYWTAEVNEEGD 600
    |||||

QY 601 TVVLSLKKFLKQDPTDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGWKGGFLLPVL 660
    |||||
Db 601 TVVLSLKKFLKQDPTDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGWKGGFLLPVL 660
    |||||

QY 661 GAVLALLFLLVLLVLLVLRKKRKIKPEPLLPEDDTRDNVFFYEGEGGEDDQDYDTQLHR 720
    |||||
Db 661 GAVLALLFLLVLLVLLVLRKKRKIKPEPLLPEDDTRDNVFFYEGEGGEDDQDYDTQLHR 720
    |||||

QY 721 GLEARPEVLRNDVAPTIIPTPMYRPRANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
    |||||
Db 721 GLEARPEVLRNDVAPTIIPTPMYRPRANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
    |||||

QY 781 DYEGSGSDAASLSLSTSSASDQDQDYDYLNEWGSFRFKKLADMYGGEDD 829
    |||||
Db 781 DYEGSGSDAASLSLSTSSASDQDQDYDYLNEWGSFRFKKLADMYGGEDD 829
    |||||
```

RESULT 14

US-10-696-639-39

```
; Sequence 39, Application US/10696639
; Publication No. US20050037439A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia Corporation
; APPLICANT: Bourner, Maureen J.
; TITLE OF INVENTION: DIFFERENTIALY EXPRESSED GENES INVOLVED IN CANCER, THE
; TITLE OF INVENTION: POLYPEPTIDES ENCODED THEREBY, AND METHODS OF USING THE SAME
; FILE REFERENCE: 01040/1
; CURRENT APPLICATION NUMBER: US/10/696,639
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: 60/422,176
; PRIOR FILING DATE: 2002-10-29
; NUMBER OF SEQ ID NOS: 3114
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 829
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-696-639-39

Query Match      100.0%; Score 4369; DB 5; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MGLPRGPLASLLLLQVLCWQCAASEPCRAVFREAEVTLKAGGAEQEPQALGKVFMGCPG 60
Db      1 MGLPRGPLASLLLLQVLCWQCAASEPCRAVFREAEVTLKAGGAEQEPQALGKVFMGCPG 60

Qy      61 QEPALFSTDDNDFTVRNGETVOERRSLKERNPLKIPFSKRILRRHKRDVWVAPISVPENG 120
Db      61 QEPALFSTDDNDFTVRNGETVOERRSLKERNPLKIPFSKRILRRHKRDVWVAPISVPENG 120

Qy      121 KGPPFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREIAK 180
Db      121 KGPPFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREIAK 180

Qy      181 YELFGHAVSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVTAT 240
Db      181 YELFGHAVSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVTAT 240

Qy      241 DEDDAIYTYNGVAVSIHQSPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
Db      241 DEDDAIYTYNGVAVSIHQSPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300

Qy      301 TDMGDGSTTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVQRLTVDLDAENSP 360
Db      301 TDMGDGSTTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVQRLTVDLDAENSP 360

Qy      361 AWRATYLLMGDDGDGHFTITTHPSNQGLITTRKGLDPEAKNQHTLYVEVTNEAPFVLKL 420
Db      361 AWRATYLLMGDDGDGHFTITTHPSNQGLITTRKGLDPEAKNQHTLYVEVTNEAPFVLKL 420

Qy      421 PTSTATIIVHVEDVNEAPVFVPPSKVVEVQSGIPTGEPVCVYTAEDPDKENQKISYRILR 480
Db      421 PTSTATIIVHVEDVNEAPVFVPPSKVVEVQSGIPTGEPVCVYTAEDPDKENQKISYRILR 480

Qy      481 DPAGHLAMPDPSGQVAVGTLDREDEQVRNNIYEVWVLAMDNQSPPTTGTGTLILLTILID 540
Db      481 DPAGHLAMPDPSGQVAVGTLDREDEQVRNNIYEVWVLAMDNQSPPTTGTGTLILLTILID 540

Qy      541 VNDHGPVPEPQITICNOSPVRHVNLITDKLSPHTSPPQALTDSDSIYWTAEVNEEGD 600
Db      541 VNDHGPVPEPQITICNOSPVRHVNLITDKLSPHTSPPQALTDSDSIYWTAEVNEEGD 600

Qy      601 TVVLSLKKFLKQDQTYDVHLSLSHGNKEQLTVIRATVCDCHGVETCPGPKWGGFILPVL 660
Db      601 TVVLSLKKFLKQDQTYDVHLSLSHGNKEQLTVIRATVCDCHGVETCPGPKWGGFILPVL 660

Qy      661 GAVTALLFLLVLLVLRKKIKIPEPLLIPEDDTRDNVFPYGGEGGEDQDYITQLHR 720
Db      661 GAVTALLFLLVLLVLRKKIKIPEPLLIPEDDTRDNVFPYGGEGGEDQDYITQLHR 720

Qy      721 GLEARPEVVLNRDVAPTIIPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780
```

```
Db      721 GLEARPEVVLNRDVAPTIIPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTLLVF 780

Qy      781 DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGEDD 829
Db      781 DYEGSGSDAASLSLTSSASDQDDYDYLNEWGSRFKKLADMYGGGEDD 829

RESULT 15
US-10-948-518-145
; Sequence 145, Application US/10948518
; Publication No. US20050064492A1
; GENERAL INFORMATION:
; APPLICANT: FREDERIC J. DESAUVAGE
; APPLICANT: GRETCHEN FRANTZ
; APPLICANT: KENNETH J. HILLAN
; APPLICANT: PAUL POLAKIS
; APPLICANT: ANDREW POLSON
; APPLICANT: VICTORIA SMITH
; APPLICANT: SUSAN D. SPENCER
; APPLICANT: THOMAS D. WU
; APPLICANT: ZEMIN ZHANG
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5026R1-US
; CURRENT APPLICATION NUMBER: US/10/948,518
; CURRENT FILING DATE: 2004-09-22
; PRIOR APPLICATION NUMBER: US/10/643,795
; PRIOR FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US 60/413,192
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US 60/419,008
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/426,847
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/484,959
; PRIOR FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 145
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-948-518-145

Query Match      100.0%; Score 4369; DB 5; Length 829;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MGLPRGPLASLLLLQVLCWQCAASEPCRAVFREAEVTLKAGGAEQEPQALGKVFMGCPG 60
Db      1 MGLPRGPLASLLLLQVLCWQCAASEPCRAVFREAEVTLKAGGAEQEPQALGKVFMGCPG 60

Qy      61 QEPALFSTDDNDFTVRNGETVOERRSLKERNPLKIPFSKRILRRHKRDVWVAPISVPENG 120
Db      61 QEPALFSTDDNDFTVRNGETVOERRSLKERNPLKIPFSKRILRRHKRDVWVAPISVPENG 120

Qy      121 KGPPFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREIAK 180
Db      121 KGPPFQRLNQLKSNKORDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREIAK 180

Qy      181 YELFGHAVSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVTAT 240
Db      181 YELFGHAVSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVTAT 240

Qy      241 DEDDAIYTYNGVAVSIHQSPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300
Db      241 DEDDAIYTYNGVAVSIHQSPKPDHLMFTIHRSTGTISVISGLDREKVPYTLTIOA 300

Qy      301 TDMGDGSTTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHVQRLTVDLDAENSP 360
```

[illegible]

Search completed: December 12, 2005, 09:19:52
Job time : 171 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2005, 09:02:47 ; Search time 235 Seconds
(without alignments)

2488.864 Million cell updates/sec

Title: US-10-696-639-39

Perfect score: 4369

Sequence: 1 MGLPRGLASLLLLQVCWLQ.....NEWGSRFKKLADMYGGEDD 829

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4369	100.0	829	1 CADH3_HUMAN	P22223 homo sapien
2	3564.5	81.6	822	2 Q8BSL6_MOUSE	Q8bsl6 mus musculus
3	3560	81.5	821	2 Q8BRE1_MOUSE	Q8bre1 mus musculus
4	3550.5	81.3	821	2 CADH3_MOUSE	P10287 mus musculus
5	3433.5	78.6	753	2 Q80VY6_MOUSE	Q80vy6 mus musculus
6	2573.5	58.9	732	1 CADHK_CHICK	P33145 gallus gall
7	2420	55.4	882	1 CADH1_HUMAN	P12830 homo sapien
8	2400.5	55.1	882	2 Q6RBF2_BOVIN	Q6rfb2 bos taurus
9	2400.5	54.9	901	2 Q9UII7_HUMAN	Q9uii7 homo sapien
10	2388	54.7	884	1 CADH1_MOUSE	P09803 mus musculus
11	2388	54.7	884	2 Q4KML8_MOUSE	Q4kml8 mus musculus
12	2388	54.7	886	1 CADH1_RAT	Q9r0t4 rattus norv
13	2387	54.6	882	2 Q5RAX1_PONPY	Q5rax1 pongo pygma
14	2377	54.4	884	2 Q6NTM0_XENLA	Q6ntm0 xenopus lae
15	2375	54.4	880	1 CADHP0_XENLA	P33148 xenopus lae
16	2371	54.3	884	1 CADH8_XENLA	P33152 xenopus lae
17	2323	53.2	491	1 CADH3_BOVIN	P19335 bos taurus
18	2274	52.0	887	1 CADH1_CHICK	P08641 gallus gall
19	2216.5	50.7	821	2 Q9UII8_HUMAN	Q9uii8 homo sapien
20	2198	50.3	872	1 CADH1_XENLA	P30344 xenopus lae
21	1893	43.3	864	2 Q90Z37_BRARE	Q90z37 brachydanio
22	1869	42.8	776	2 Q4TX7_TETNG	Q4tx7 tetraodon n
23	1715	39.3	906	2 Q5RAX1_PONPY	Q5rax1 pongo pygma
24	1714	39.2	906	1 CADH2_HUMAN	P19022 homo sapien
25	1704.5	39.0	912	1 CADH2_CHICK	P10288 gallus gall
26	1702	39.0	906	2 Q8N173_HUMAN	Q8n173 homo sapien
27	1701	38.9	906	2 Q8BSI9_MOUSE	Q8bsi9 mus musculus
28	1697.5	38.9	906	2 Q6GU11_MOUSE	Q6gu11 mus musculus
29	1695	38.8	1101	2 Q4TX6_TETNG	P15116 mus musculus
30	1695	38.8	1101	2 Q4TX6_TETNG	Q4tx6 tetraodon n
31	1694.5	38.8	875	2 Q5RB18_PONPY	Q5rb18 pongo pygma

RESULT 1

ID	CADH3_HUMAN	STANDARD;	PRT;	829 AA.
AC	P22223;			
DT	01-AUG-1991 (Rel. 19, Created)			
DT	01-AUG-1991 (Rel. 19, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DE	Cadherin-3 precursor (Placental-cadherin) (P-cadherin).			
GN	Name=CDH3; Synonyms=CDHP;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]_TaxID=9606;			
RP	NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=90090501; PubMed=2793940; DOI=10.1093/jcb.109.4.1787;			
RA	Shimoyama Y., Yoshida T., Terada M., Shimamoto Y., Abe O.,			
RA	Hirohashi S.;			
RT	"Molecular cloning of a human Ca2+-dependent cell-cell adhesion			
RT	molecule homologous to mouse placental cadherin: its low expression in			
RT	human placental tissues";			
RL	J. Cell Biol. 109:1787-1794 (1989).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].			
RC	TISSUE=Testis;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klauser R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,			
RA	Schmerch A., Schein J.E., Jones S.J.M., Marta M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE OF 1-15.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=99111163; PubMed=9815605;			
RA	Jarrard D.F., Paul R., Van Bokhoven A., Nguyen S.H., Bova G.S.,			
RA	Wheelock M.J., Johnson K.R., Schaiken J.J., Bussemakers M., Isaacs W.B.;			
RT	"P-cadherin is a basal cell-specific epithelial marker that is not			

P24503 gallus gall
Q92ly3 rattus norv
P19534 bos taurus
Q90275 brachydanio
Q4rvh8 tetraodon n
P39038 mus musculus
Q9b205 homo sapien
Q8nb64 homo sapien
P55283 homo sapien
P33147 xenopus lae
P20310 xenopus lae
Q4uli6 brachydanio
Q4836 tetraodon n
P79883 xenopus. ma

RT expressed in prostate cancer.";
RL Clin. Cancer Res. 3:2121-2128(1997).
RN [4].
RP TISSUE SPECIFICITY.
RX MEDLINE=89195140; PubMed=2702654;
RA Shimoyama Y., Hirohashi S., Hirano S., Noguchi M., Shimoseato Y.,
RA Takeichi M., Abe O.;
RT "Cadherin cell-adhesion molecules in human epithelial tissues and
RT carcinomas";
RL Cancer Res. 49:2128-2133(1989).
RN [5].
RP DISEASE.
RX MEDLINE=21470328; PubMed=11544476; DOI=10.1038/ng716;
RA Sprecher E., Bergman R., Richard G., Lurie R., Shalev S.,
RA Petronius D., Shalata A., Ambinder Y., Leibin R., Perlman I., Cohen N.,
RA Szargel R.;
RT "Hypotrichosis with juvenile macular dystrophy is caused by a mutation
RT in CDH3, encoding P-cadherin";
RL Nat. Genet. 29:134-136(2001).
RN [6].
RP VARIANT HJMD HIS-503.
RX MEDLINE=22336734; PubMed=12445216;
RX DOI=10.1046/j.1523-1747.2002.19528.x;
RA Indelman M., Bergman R., Lurie R., Richard G., Miller B.,
RA Petronius D., Ciubutaro D., Leibin R., Sprecher E.;
RT "A missense mutation in CDH3, encoding P-cadherin, causes
RT hypotrichosis with juvenile macular dystrophy.";
RL J. Invest. Dermatol. 119:1210-1213(2002).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in some normal epithelial tissues
CC and in some carcinoma cell lines.
CC -!- DISEASE: Defects in CDH3 are the cause of hypotrichosis with
CC juvenile macular dystrophy (HJMD) [MIM:601553]. HJMD is a rare
CC autosomal recessive disorder characterized by early hair loss
CC heralding severe degenerative changes of the retinal macula and
CC culminating in blindness during the second to third decade of
CC life.
CC -!- SIMILARITY: Contains 5 cadherin domains.
CC -!- DATABASE: NAME=Mutations of the CDH3 gene;
CC NOTE=Retina International's Scientific Newsletter;
CC WWW="http://www.retina-international.com/sci-news/cdh3mut.htm".
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X63629; CAA45177.1; -; mRNA.
DR EMBL; BC041846; AAA41846.1; -; mRNA.
DR EMBL; X95824; CAA65093.1; -; Genomic_DNA.
DR PIR; A33659; IJHUCP.
DR HSP; P09803; 117W.
DR SMR; P22223; 749-824.
DR Ensembl; ENSG00000062038; Homo sapiens.
DR HGNC; HGNC:1762; CDH3.
DR MIM; 114021; --.
DR MIM; 601553; --.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR00126; Cadherin.
DR InterPro; IPR000233; Cadherin_C-term.
DR Pfam; PF00028; Cadherin; 5.
DR Pfam; PF01049; Cadherin C; 1.
DR PRINTS; PR0205; CADHERIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS0268; CADHERIN_2; 5.
DR Calcium; Cell adhesion; Disease mutation; Glycoprotein; Repeat;
KW Sensory transduction; Signal; Transmembrane; Vision.

	FT	SIGNAL	1	24	Potential.
FT	PROPEP	25	107		Cadherin-3.
FT	CHAIN	108	829		Extracellular (Potential).
FT	TOPO_DOM	108	654		Potential.
FT	TRANSMEM	655	677		Cytoplasmic (Potential).
FT	TOPO_DOM	678	829		Cadherin 1.
FT	DOMAIN	108	215		Cadherin 2.
FT	DOMAIN	216	328		Cadherin 3.
FT	DOMAIN	329	440		Cadherin 4.
FT	DOMAIN	441	546		Cadherin 5.
FT	DOMAIN	547	650		Ser-rich.
FT	COMBIAS	785	800		N-linked (GLNAC. . .) (Potential).
FT	CARBOHYD	200	200		N-linked (GLNAC. . .) (Potential).
FT	CARBOHYD	566	566		R -> H (in HJMD).
FT	VARIANT	503	503		/FTId=VAR_015422.
FT	CONFLICT	237	237		V -> M (in Ref. 2).
FT	CONFLICT	563	563		H -> Q (in Ref. 2).
FT	SEQUENCE	829	AA; 91427	MW; E503CFFFE5D981F1	CRC64;
Query Match			100.0%;	Score 4369;	DB 1; Length 829;
Best Local Similarity			100.0%;	Pred. No. 2.3e-264;	
Matches 829;			Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	MGLPRGLASLLLLQVLCQAASEPCRAVFREAEVTLAEAGAEQEPGQALGVFMGCPG	60		
Db	1	MGLPRGLASLLLLQVLCQAASEPCRAVFREAEVTLAEAGAEQEPGQALGVFMGCPG	60		
Qy	61	QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKILRRHKRDWVAPISVPENG	120		
Db	61	QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKILRRHKRDWVAPISVPENG	120		
Qy	121	KGPPQRLNOLKSNKORDTKIFYSITGPGADSPGEGVFAVEKETGMLLNKPLDREIAK	180		
Db	121	KGPPQRLNOLKSNKORDTKIFYSITGPGADSPGEGVFAVEKETGMLLNKPLDREIAK	180		
Qy	181	YELFGHAVSENGASVEDPMNISIIVTDQNDHKPKFTQDTPRGSVLEGVLPGTSVMQVTAT	240		
Db	181	YELFGHAVSENGASVEDPMNISIIVTDQNDHKPKFTQDTPRGSVLEGVLPGTSVMQVTAT	240		
Qy	241	DEDDAIYTYNGVAVSIHSEPKDPDHLMTIHRSTGTISVISGGLDREKVPETLTIQA	300		
Db	241	DEDDAIYTYNGVAVSIHSEPKDPDHLMTIHRSTGTISVISGGLDREKVPETLTIQA	300		
Qy	301	TDMDGSGSTTTAVAVVEILDANDNPNFDPQKEAHVPENAVGHEVORLTVTDLDAENSP	360		
Db	301	TDMDGSGSTTTAVAVVEILDANDNPNFDPQKEAHVPENAVGHEVORLTVTDLDAENSP	360		
Qy	361	AMRATYILMGDDGDHFTITTHPESNQGLITTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420		
Db	361	AMRATYILMGDDGDHFTITTHPESNQGLITTRKGLDFAKNQHTLYVEVTNEAPFVLKL	420		
Qy	421	PTSTATIVHVEDVNEAPVFPVPPSKVVEQEGIPTEGPPCVVYTAEDPKENQKISYILR	480		
Db	421	PTSTATIVHVEDVNEAPVFPVPPSKVVEQEGIPTEGPPCVVYTAEDPKENQKISYILR	480		
Qy	481	DPAGWLAMPDSCGVTAAGTLDREDFQVFNNNIYEVNVLAWDNGSPPTTGTGTLTLLID	540		
Db	481	DPAGWLAMPDSCGVTAAGTLDREDFQVFNNNIYEVNVLAWDNGSPPTTGTGTLTLLID	540		
Qy	541	VNDHGPVPEPRQITICNQSPVRHVLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD	600		
Db	541	VNDHGPVPEPRQITICNQSPVRHVLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD	600		
Qy	601	TVVLSLKKFLKQDTPYVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGFILPVL	660		
Db	601	TVVLSLKKFLKQDTPYVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGFILPVL	660		
Qy	661	GAVALILFLLVLLVLLVLRKKIKIPELLPDDTRDNVYFYGGEGGEEQDDYDITQHR	720		
Db	661	GAVALILFLLVLLVLLVLRKKIKIPELLPDDTRDNVYFYGGEGGEEQDDYDITQHR	720		
Qy	721	GLEARPEVLRNDVAPTIIPTPMYRPRPANPDEIGNFIIENLKAANTDPTAPPYDTLLVF	780		

Db 721 GLEAREVLRNDVATPTPTMYRPRPANPDBEIGNFIENLKAANTDTPAPPYDTLVLF 780
QY 781 DYEGSGSDAASLSSTSSADODQDYDLNWSGRFKKLADMYGGGEDD 829
Db 781 DYEGSGSDAASLSSTSSADODQDYDLNWSGRFKKLADMYGGGEDD 829
RESULT 2
Q8BSL6 MOUSE
ID Q8BSL6_MOUSE PRELIMINARY; PRT; 822 AA.
AC Q8BSL6_MOUSE PRELIMINARY; PRT; 822 AA.
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 13-SEP-2005 (TREMELrel. 31, Last annotation update)
DE Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length enriched
DE library, clone:5930439E13 product:cadherin 3, full insert sequence
DE (Cadherin 3).
DE Names:Cdh3;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadofa K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [5]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX ADACHI J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Haehizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa T., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RN Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong E.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RG NIH MGC Project;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK031265; BAC27327.1; -; mRNA.
DR EMBL; BC098459; RAH98459.1; -; mRNA.
DR HSSP; P09803; 117X.
DR Ensemble; ENSMUSG00000061048; Mus musculus.
DR MGI; MGI:88356; Cdh3.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; IDA.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0016337; P:cell-cell adhesion; IMP.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C term.
DR Pfam; PF01028; Cadherin; 5.
DR Pfam; PF01049; Cadherin; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 3.

enriched library, clone: B130021C20 product: cadherin 3, full insert sequence.

DE Name=Cdh3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Arakawa T., Hara A., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Hayashaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
RA The FANTOM Consortium,
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
RX MEDLINE=20493374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaehiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaehiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaehiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaehiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaehiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaehiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaehiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaehiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;

Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishi Y., Itoh M., Kagawa I., Kasukawa T., Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito K., Saichon K., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akaira S., Takada Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.; Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK045041; BAC32194.1; -, mRNA.

DR HSSP; P09803; 117X.

DR MGI; MGI:88356; Cdh3.

DR GO; GO:0005615; C:extracellular space; TAS.

DR GO; GO:0016021; C:integral to membrane; TAS.

DR GO; GO:0005886; C:plasma membrane; IDA.

DR GO; GO:0016337; P:cell-cell adhesion; IMP.

DR InterPro; IPR002126; Cadherin.

DR InterPro; IPR000233; Cadherin_C_term.

DR Pfam; PF00028; Cadherin; 5.

DR Pfam; PF01049; Cadherin_C; 1.

DR PRINTS; PR00205; CADHERIN.

DR SMART; SM00112; CA; 4.

DR PROSITE; PS00232; CADHERIN_1; 3.

DR PROSITE; PS00268; CADHERIN_2; 4.

KW Calcium.

SQ SEQUENCE 821 AA; 90513 MW; 168B56C5FB5CCED0 CRC64;

Query Match 81.5%; Score 3560; DB 2; Length 821;

Best Local Similarity 81.6%; Pred. No. 8.9e-214;

Matches 679; Conservative 56; Mismatches 83; Indels 14; Gaps 5;

QY 1 MGIPRGLA-SLLQLQVCMLOCAASPCRAVF-REAETVLEAGSAQEPQALGKVFMC 58

Db 1 MEULSGPHALLLQLQVCMVVSVSPYRAGFTGEAGVILEVETGDLQPSQVLGKALAG-- 59

QY 59 PGQEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKRLRRHKRDVVVAPISVPE 118

Db 60 -----QGMHADNGDIIMLTGTGQGGKAMHS-----PPTILRRKREWMPPIFVPE 109

QY 119 NGKGPPFQRLNQLKSNKDRDKTIFYITGPGADSPGEGFVAVKETGWLNLNKLPLDREI 178

Db 110 NGKGPPFQRLNQLKSNKDRDKTIFYITGPGADSPGEGFVAVKETGWLNLNKLPLDREI 169

QY 179 AKYELCHAVSENGASVEDPMNISIIVTDNDHKPKFTQDTFGSVLEGLVPGTSVMQVT 238

Db 170 VKYELCHAVSENGASVEEPMNISIIVTDQNDNKPFTQDTFGSVLEGLVPGTSVMQVT 229

QY 239 ATDEDDAIYNGVWYSHSQPKDPKPHDLMTFTHRSTGTISVISGLDREKVPYTLTI 298

Db 230 ATDEDDAVNTYNGVWYSHSQPKDPKPHDLMTFTHRSTGTISVISGLDREKVPYTLTI 289

QY 299 QATDMGDSGTTTAVAVVEILDANDNAPMDPPQKYEAHVPENAVGHEVQRLTVDLDAPN 358

Db 290 QATDMGDSGTTTAAVAVVQILDANDNAPPEPQKYEAWPENEVEGHEVQRLTVDLDVFN 349

QY 359 SPAWRATYLIMGDDGDFHTITTHPSNQILTRKGLDPEAKNQHTLVYEVNTEAPFVL 418

Db 350 SPAWRATYHIVGGDGDHFTITTHPETNQGLVTKKGLDPEAQDQHTLVYEVNTEAPFAV 409

QY 419 KLPTSTATIVHVEDVNEAPFVPPSKVVEVQSGIPTGPEVCVYTAEDPDKENOKISYRI 478

Db 410 KLPTATATVHVHVDVNEAPFVPPSKVIEAQSGISIGELVCIYTAQDPDKDQKISYTI 469

QY 479 LRDPAGWLAMPDPSGQVAVTGLDREDEQFVRNNIYEVWVLMNDNGSPPTTGTGTLTLL 538

Db 470 SRDPANWLAVDPSGQITAAAGILDRDEQFVKNNVYEVWVLMATDSGNPPTTGTGTLTLL 529

QY 539 IDVNDHGVPPEPQITICNOSPVRVNLNITDKLSPTSPFOQLTDDSDIYTAEVNEE 598

Db 530 TDINDHGPIPEPQIIICNOSPVPQVNLNITDKLSNPSPPFOQLTHDSDIYMAEVSEK 589

QY 599 GDTVVLSLKKFLKQDYDVHLSLSDHGKKEQLTVIRATVCDCHGHV-ETCPGPKGGFTL 657

Db 590 GSTVALSLKKFLKQDYDVHLSLSDHGKKEQLTVIRATVCDCHGHV-FNDCPRPKGGFTL 649

QY 658 PVLGAVLALLFLLVLLVLLVLRKKIKKEPLLPEDDTRDNVFFYGEGBEGDQDYDITQ 717

Db 650 FLGAVLALLTLLVLLVLLVLRKKIKKEPLLPEDDTRDNVFFYGEGBEGDQDYDITQ 709

QY 718 LHRGLEARPEVVLNRDVAPTIITPMYRPRNPANPEIGNFIENLKAANTDPTAPPYDTL 777

Db 710 LHRGLEARPEVVLNRDVAPTIITPMYRPRNPANPEIGNFIENLKAANTDPTAPPYDSL 769

QY 778 LVFDYEGSGDAASSLTSSASDQDQDYDYLNWGSRKFKLADMYGGGDD 829

Db 770 LVFDYEGSGDAASSLTSSASDQDQDYDYLNWGSRKFKLADMYGGGDD 821

RESULT 4

CADH3_MOUSE

ID CADH3_MOUSE STANDARD; PRT; 822 AA.

AC P10287; Q61465;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Cadherin-3 precursor (Placental-cadherin) (P-cadherin).

GN Name=Cdh3; Synonyms=Cdhp;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muroidae; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=88111554; PubMed=3428270;

RA Nose A., Nagafuchi A., Takeichi M.;

RT "Isolation of placental cadherin cDNA: identification of a novel gene family of cell-cell adhesion molecules.";

RL EMBO J. 6:3655-3661(1987).

RN [2]

RP PARTIAL NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6; TISSUE=Embryo;

RX MEDLINE=91360343; PubMed=1886768;

RA Hatta M., Miyatani S., Copeland N.G., Gilbert D.J., Jenkins N.A., Takeichi M.;

RT "Genomic organization and chromosomal mapping of the mouse P-cadherin gene.";

RL Nucleic Acids Res. 19:4437-4441(1991).

RN [3]

RP NUCLEOTIDE SEQUENCE OF 1-55.

RX MEDLINE=93294853; PubMed=8515462;

RA Faraldo M.L., Cano A.;

RT "The 5' flanking sequences of the mouse P-cadherin gene. Homologies to 5' sequences of the E-cadherin gene and identification of a first 215 base-pair intron.";

RL J. Mol. Biol. 231:935-941(1993).

RN [4]

RP NUCLEOTIDE SEQUENCE OF 1-55.

RC STRAIN=C57BL/6;

RA Hatta M., Takeichi M.;

RT "Complex cell type-specific transcriptional regulation by the promoter and an intron of the mouse P-cadherin gene.";

RL Dev. Growth Differ. 36:509-519(1994).

RN [5]

RP DEVELOPMENTAL STAGE.

RC STRAIN=C57BL/6; TISSUE=Testis;

RX MEDLINE=97033837; PubMed=8879495;

RA Munro S.B., Blaschuk O.W.;

RT "A comprehensive survey of the cadherins expressed in the testes of fetal, immature, and adult mice utilizing the polymerase chain reaction.";

RL Biol. Reprod. 55:822-827(1996).

CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC052189; AAH52189.1; -, mRNA.
DR HSSP; P09803; 117X.
DR MGI; MGI:88356; Cdh3.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0016337; P:cell-cell adhesion; IMP.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C-term.
DR Pfam; PF00028; Cadherin; 5.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS0268; CADHERIN_2; 4.
KW Calcium.
FT NON_TER.
SQ SEQUENCE 753 AA; 8330 MW; 414AFB4059CE94D2 CRC64;

Query Match 78.6%; Score 3433.5; DB 2; Length 753;
Best Local Similarity 87.1%; Pred. No. 6.4e-206;
Matches 639; Conservative 51; Mismatches 43; Indels 1; Gaps 1;

QY 97 PSKRILRRHRKDWVAPISVPENGKGFPPQRLNQLSKNDRDKPKIYISITGPGADSPPEG 156
DB 20 PPRILRRKREWMPEIPVPENGKGFPPQRLNQLSKNDRDKPKIYISITGPGADSPPEG 79

QY 157 VFAVEKETGMLLNKPLDREIAKYELFGHAVSENGASVEDPMNISIIIVTDQNDHKPKPT 216
DB 80 VFTIEKESGWLHMLPLDREIKVYELYGHAVSENGASVEEPMNISIIIVTDQNDHKPKPT 139

QY 217 QDTFRSGVLEGLVPGTSVMQVATDEDDALTYTNGVAVSIHQEPKDPHDLMTTHRT 276
DB 140 QDTFRSGVLEGLVPGTSVMQVATDEDDAVNTYNGVAVSIHQEPKDPHDLMTTHRT 199

QY 277 GTISVTSGLDRKVPREYLTICATDMGDSSTTAVVVEILDANDNAPMFPDQKYEAH 336
DB 200 GTISVTSGLDRKVPREYLTICATDMGDSSTTAVVVEILDANDNAPMFPDQKYEAW 259

QY 337 VPENAVGHEVQRLTVTDLDAPNSPAWRATYILMGDDGDHFTITTHPESNQGLTTRKGL 396
DB 260 VPENAVGHEVQRLTVTDLDVPSNPAWRATYIIVGGDDGDHFTITTHPESNQGLTTRKGL 319

QY 397 DFRAKQHTLYVEVNEAPVLPSTATIVVHVEDVNEAPVFPVPSKVEVQEGIPTG 456
DB 320 DFRAKQHTLYVEVNEAPVLPSTATATVTVVHKVDNEAPVFPVPSKVEVQEGISIG 379

QY 457 EPVCVVTAEPPDKENOKISVRIILRDPAGVLAMPDPSQVTAAGTLDREDEQFVRNNIYEV 516
DB 380 ELVCITYTAQPDKEQDKISITISRDPAWLAADPDGQITAAIGLREDEQFVKNVYEV 439

QY 517 MVLAMNGSPPTTGTGTLLTLTDVNDHGVPPFPRTIICNQSPFVRHVLNITDKLSPT 576
DB 440 MVLATSGNPPTTGTGTLLTLTDINDHGPPIPRQIICNQSPVQVNLITDKLSPTS 499

QY 577 SPFQAQLTDSDIYWTAEVNEEGDVTLSLKKFLKQDVTYVHLSHDGNGKEQLTVIRAT 636
DB 500 SPFQAQLTDSDIYWMAEVSEKGTVALSLKKFLKQDVTYVHLSHDGNGKEQLTVIRAT 559

QY 637 VCDCHGV-ETCPGPKWGGFIPVLGAVLALLFLLVLLVLLVKKKIKKEPLLPPEDTR 695
DB 560 VCDCHGVQVFNDCPRPKWGGFIPVLGAVLALLFLLVLLVLLVKKKIKKEPLLPPEDTR 619
QY 696 DNIFYEGEGGEDQDITQLHGLEARPEVLRNDVAPTIPTPMYRPRANPDEIG 755

DB 620 DNIFYEGEGGEDQDITQLHGLEARPEVLRNDVAPTIPTPMYRPRANPDEIG 679
QY 756 NFIIENLKAANTDPTAPPYDTLLVFDYEGSGDAASLSLTSSASDQDQDYDLNWSGR 815
DB 680 NFIIENLKAANTDPTAPPYDTLLVFDYEGSGDAASLSLTSSASDQDQDYDLNWSGR 739
QY 816 FKLLADMYGGGEDD 829
DB 740 FKLLADMYGGGEDD 753

RESULT 6
CADHK_CHICK
ID CADHK_CHICK STANDARD; PRT; 732 AA.
AC P33145;
DC 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE B-cadherin precursor (K-CAM protein) (Fragment).
GN Name=K-CAM;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92107987; PubMed=1763068;
RA Sorkin B.C., Gallin W.J., Edelman G.M., Cunningham B.A.;
RT "Genes for two calcium-dependent cell adhesion molecules have similar
structures and are arranged in tandem in the chicken genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:11545-11549 (1991).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 7-732.
RC TISSUE=Embryonic brain;
RX MEDLINE=91225083; PubMed=2026653; DOI=10.1083/jcb.113.4.893;
RA Napolitano E.W., Venstrom K., Wheeler E.F., Reichardt L.F.;
RT "Molecular cloning and characterization of B-cadherin, a novel chick
cadherin.";
RL J. Cell Biol. 113:893-905 (1991).
CC -!- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
They preferentially interact with themselves in a homophilic
manner in connecting cells; cadherins may thus contribute to the
sorting of heterogeneous cell types. B-cadherin may have important
functions in neurogenesis, in at least some epithelia, and in
embryogenesis.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in a wide variety of tissues.
CC -!- SIMILARITY: Contains 5 cadherin domains.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC EMBL; M81894; AAA48929.1; -, Genomic_DNA.
DR EMBL; X58518; CAA41408.1; -, mRNA.
DR PIR; A41634; IJCHCB.
DR HSSP; P09803; 117X.
DR SMR; P33145; 7-224, 651-726.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C-term.
DR Pfam; PF00028; Cadherin; 5.
DR Pfam; PF01049; Cadherin_C; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS0268; CADHERIN_2; 5.
KW Calcium; Cell adhesion; Glycoprotein; Repeat; Transmembrane.
FT PROPEP <1 6 Potential.
FT CHAIN 7 732 B-cadherin.

```
FT TOPO_DOM 6 554 Extracellular (Potential).
FT TRANSMEM 555 580 Potential.
FT TOPO_DOM 581 732 Cytoplasmic (Potential).
FT DONAIN 6 114 Cadherin 1.
FT DONAIN 115 227 Cadherin 2.
FT DONAIN 228 339 Cadherin 3.
FT DONAIN 340 443 Cadherin 4.
FT DONAIN 444 554 Cadherin 5.
FT COMPBTAS 689 702 Ser-rich.
FT CARBOHYD 137 137 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 410 410 N-linked (GlcNAc. .) (Potential).
FT CONFLICT 414 414 M -> V (in Ref. 2).
FT NON_TER 1
SQ SEQUENCE 732 AA; 80613 MW; 091D59A6A16CBD45 CRC64;

Query Match 58.9%; Score 2573.5; DB 1; Length 732;
Best Local Similarity 66.2%; Pred. No. 3.6e-152;
Matches 485; Conservative 103; Mismatches 138; Indels 7; Gaps 5;

Qy 102 LRRHKRWVAVIPSPENGKGFPPQRLNQLKSKDRDTKIFYSITGPGADSPPEGVFAVE 161
Db 1 LRRQKRWVPIPKVPENEGFPFNVLQIKSNDRDAKIFYSITGQADAPPEGIETIE 60

Qy 162 KETGLLLNKPDLREEIAKELFGLHVSSENGASVEDPMNISIIIVTDQNDHKPKFTQDTR 221
Db 61 KETGMKVKVQPLDREHINKYHLHSHVSENGKPEEPMEIIVTVDQNDNKPQFTQSVFR 120

Qy 222 GSVLEGVLPSTVMQVATDEDDAIYTYNGVVAISHSQPKDPHDLMTIHRSTGISV 281
Db 121 GSVPEGALPGTSVMRVNATDADDVETNGYVIAYSILSQEPREPHEMFTVNRATGTLVS 180

Qy 282 ISSGLDRKPEYTLTQATDMCGDGGSTTTAVAVVEILDANDNAPMFDPOKYEAAHPENA 341
Db 181 IASGLDRERVREYITLWQADLDQGLTTTALAVIEITVDNDNAPEDFPKTYEAAVENE 240

Qy 342 VGHEVQRLVTDLPANSPAWRATYLYINGSDGDGHFTTTHPESNOGILTRKGLDPEAK 401
Db 241 AELEVARLATTDLDEPHTPAWRAVYSIVRGNEGGAFTITDPAFNEGVLRTAKGLDYEAK 300

Qy 402 NQHTLYEVNTAEAPVLKLPSTATIVVHVEDUNEAPVFPSPKVVVEQGIPTGEVVCV 461
Db 301 RQFLVHAVVNEAPFAIKLPTATATVMVSDVNEAPVFPDPLSLAQVPEDVPLGQPLAS 360

Qy 462 YTAEDPDK-ENQKLSYRILTRDPAGWLAMPDPSGOVTVAGTLDRDEDFVRNNIYEVMLA 520
Db 361 YTAQDPRAQQRIKYVMSGSDPAGWLAVHPENGLITAREQLDRE-SPTKNSITMAVILLA 419

Qy 521 MDNGSPPTGTGTLTLLTLDVNDHGPVPEPQITICNQSPYRVHVLNITDKDLSPTSPFQ 580
Db 420 VDDGLPPATGATGTLTLLTLDVNDHGPPEPRDIVICNRSVPVQVLITTDRLDLPNTGPF 479

Qy 581 AQLTDDSDIYTAEVNEBGDTVVLSLKFLKQDTYDVHLSDSHGKNEQLTVIRATVDCD 640
Db 480 AELSHGSGDSWAVEGVNGSDTVALWLTPELQNLYSVYVLRFLDRQGDQVTVIRAQVCD 539

Qy 641 HGHVETC---PGPWKG-GFILLPVLAGVLLALLVLLLVLRKKKIKPELLEDPTD 696
Db 540 QGRVESCAQKPRVDTGVPILVAVLAGVALLVLLLVLRKKVKEPILLLEDPTD 599

Qy 697 NVFYEGEGGEEDQDYDITQHLRGLEAPVWLNRNDVAPITIIPTMYRPRPANPDEIGN 756
Db 600 NIFYEGEGGEEDQDYDLSQLHGLDARPEVI-RNDVAPPLMAAPQVYRPRPANPDEIGN 658

Qy 757 FIENLKAANTDPTAPPYDITLLVFDYEGSGDAAASLSLTSSASDQDDYDYLNEWSRF 816
Db 659 FIDENLKAANTDPTAPPYDLSLLVFDYEGGSEATSLSLNSSASDQDDYDYLNEWGNRF 718

Qy 817 KKLADMYGGGEDD 829
Db 719 KKLAELYGGGEDE 731
```

RESULT 7

```
CADH1_HUMAN
ID CADH1_HUMAN STANDARD; PRT; 882 AA.
AC P12830; Q13799; Q14216; Q15855; Q16194; Q4PJ14;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Epithelial-cadherin precursor (E-cadherin) (Uvomorulin) (Cadherin-1) (CAM 120/80).
GN Name=CDH1; Synonyms=CDHE, UVO;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Pancreas;
RX MEDLINE=93211394; PubMed=8459805;
RA Bussemakers M.J.G., Mees S.G.M., van Bokhoven A., Debruyne F.M.J.,
RA Schaiken J.A.;
RT "Molecular cloning and characterization of the human E-cadherin
RT cDNA.";
RL Mol. Biol. Rep. 17:123-128 (1993).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Liver;
RX MEDLINE=94242050; PubMed=8185635;
RA Rimm D.L., Morrow J.S.;
RT "Molecular cloning of human E-cadherin suggests a novel subdivision of
RT the cadherin superfamily.";
RL Biochem. Biophys. Res. Commun. 200:1754-1761 (1994).
RN [4]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS PRO-478; THR-617;
RP MET-832 AND LYS-880.
RA Livingston R.J., Rieder M.J., Shaffer T., Bertucci C., Baier C.N.,
RA Rajkumar N., Willa H.T., Daniels M., Downing T.K., Stanaway I.B.,
RA Nguyen C.P., Gildersleeve H., Cassidy C.M., Johnson E.J.,
RA Swanson J.E., McFarland I., Yool B., Park C., Nickerson D.A.;
RT "NIHES-SNPs, environmental genome project, NIHES ES15478, Department
RT of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE OF 56-882.
RC TISSUE=Placenta;
RX MEDLINE=95324920; PubMed=7601454;
RA Berx G., Staes K., van Hengel J., Molenans F., Bussemakers M.J.G.,
RA van Bokhoven A., van Roy F.;
RT "Cloning and characterization of the human invasion suppressor gene E-
RT cadherin (CDH1).";
RL Genomics 26:281-289 (1995).
RN [6]
RP NUCLEOTIDE SEQUENCE OF 172-311.
RC TISSUE=Liver;
RX MEDLINE=89031725; PubMed=3263290;
RA Mansouri A., Spurr N., Goodfellow P.N., Kemler R.;
RT "Characterization and chromosomal localization of the gene encoding
RT the human cell adhesion molecule uvomorulin.";
RL Differentiation 38:67-71 (1988).
RN [7]
RP NUCLEOTIDE SEQUENCE OF 265-392.
RC TISSUE=Liver;
RA Frixen U.H.;
RL Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.
RN [8]
RP NUCLEOTIDE SEQUENCE OF 1-16.
RX MEDLINE=94380041; PubMed=8093045;
RA Bussemakers M.J.G., Girolodi L.A., van Bokhoven A., Schaiken J.A.;
RT "Transcriptional regulation of the human E-cadherin gene in human
```



```
Db 188 VFSITGQADTPVGVFIIRERETGWLKVTEPLDRERIATVTLFSAVSSNGNAVEDPME 247
Qy 201 ISIIITDNDHKPKFTQDTFRGSLVGLPGTSMQVTADEDDAIYTYNGVAYSHSQ 260
Db 248 ILITVTDNDNDKPFTEQVFKGSMEGALPGTSMVETATDADDDVNTYNAATYTLSS 307
Qy 261 EPKDPHDLMTIHRSTGTISVISGLDREKVEPYTLTIQATDMDGSGSTTAVAVBILD 320
Db 308 DPFLPKNMFTINRNTGVISVTTGLDRESFPTLVVQAADLOGEGLSITATAVITVD 367
Qy 321 ANDNAPMFDPOKYBAHPVNAVGHVQRLVTDLDAPNSPAWRATYILIMGGDDGDHFTIT 380
Db 368 TNDNPPFNPPTYKGVQVPEANVVIITLTKVDADAPNTPAWEAVYIIL-NDGQGFVVT 426
Qy 381 THPESNOGILTRKGLDFEAKNQHTLYVEVNEAPFVLKPTSTATIVHVEDVNEAPVF 440
Db 427 TNPVNDGILKTAKGLDFEAKQQYILHVAVTNVVPEVSLTSTATVTVDVLVNEAPIF 486
Qy 441 VPPSKVVEVOEGIPTGEPVCVYTAEDPK-ENQKISYRIILRDPAWLMADPDSQGVTAAG 499
Db 487 VPPEKRVEVSEDFVGQEIYSYTAQEDTFMEQKITVRIWRDTANWLEINPDTGAISTRA 546
Qy 500 TLDREDFQVRNNIYEVNVLAMONGSPPTTGTGTLTLTLIDVNDHGPVPEPRQITICNS 559
Db 547 ELDRDPEHVKNSTYALIIATDNGSPVATGTGLLILSDVNDNAPIPEPRTIFFCERN 606
Qy 560 PVRHVLNITDKLSPHTSPPQAQLTDDSDIYWTAEVNE-EGDVTVLSKKFLKQDITDVH 618
Db 607 PKQVINIADLPNTPSPETAELTHGASANWTIYNDPTQESIIILPKMALEVGYKIN 666
Qy 619 LSLSDHNGKEQLTVIRATVCDCHGVETC--PGPWKGGF---LPLVGLAVLALLFLLV 672
Db 667 LKLMNDQNKQDVTTLEVSVCDCGAGVCRKAQVFEAGLQIPALGLIGLILALLILL 726
Qy 673 LLLVLRKKRIKEPLLLPEDDTRDNVYFYEGEGGEDQDYDITQLHRGLEARPEVVLRN 732
Db 727 LLLFLRRRAVVKPELLPEDDTRDNVYFYEGEGGEDQDQLSGLHGLDARPEVT-RN 785
Qy 733 DVARTIIPMYRPRRPNPDBIGNFIENLKAANTDPTAPPYDILLVFDYEGSGSDAASL 792
Db 786 DVAPTLMSVPRYLRPNPDBIGNFIDENLKAADTPTAPPYDLSLLVFDYEGSGSEASL 845
Qy 793 SSLTSSASDQDDYDYLNEGWSRPFKLADMYGGGEDD 829
Db 846 SSLNSESSEDDQDDYDYLNEGWRPFKLADMYGGGEDD 882

RESULT 8
Q6R8F2_BOVIN PRELIMINARY; PRT; 882 AA.
ID Q6R8F2_BOVIN PRELIMINARY; PRT; 882 AA.
AC Q6R8F2;
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TremBLrel. 27, Last annotation update)
DE E-cadherin.
GN Name=CDH1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hanna C.B., Cox M.L., Golding M.C., Westhuesin M.E., Kraemer D.C.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY508164; AAR91598.1; -; mRNA.
DR SRR; O6R8F2; 155-372, 802-877.
DR CO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005059; F:calcium ion binding; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C term.
```

```
DR Pfam; PF00028; Cadherin; 5.
DR Pfam; PF01049; Cadherin_C; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS02268; CADHERIN_2; 4.
KW Calcium.
SQ SEQUENCE 882 AA; 97938 MW; 0EC5C82ADF7D5F36 CRC64;

Query Match 55.1%; Score 2406.5; DB 2; Length 882;
Best Local Similarity 53.8%; Pred. No. 1.3e-141;
Matches 476; Conservative 120; Mismatches 225; Indels 63; Gaps 12;

Qy 6 GPLA-SLLLIQVC-----WLOCAASBPICRAVFEAEVTEAGGAEOBPQALGKV-FMGQ 58
Db 2 GPNRSLSALCCCRCPNWL-CREPEPCIPGFAESYTFVPRNLERGRVILGRVSEGC 60
Qy 59 PGQEPALFSTDNDFTVRNGETVQERRSLKERNP----- 92
Db 61 AGLPRTVYVSDTRFKVHTDGLTVRRPVHLHRLPELSFLVHAWDSTHRKLSKTVLEVSA 120
Qy 93 -----LKIIFSKRI-LRRHKRDVAVAPISVPENKGGPPQRLNOLKS 133
Db 121 HHHHHSHDPSGTOFTEVLTFFGPHHGLRRQRDWIPIPCPENKGGPPKSLVOIKS 180
Qy 134 NKDRDTKIFYSITPGADSPPEGVFAVEKETGMLLNKPLDRBEIAKYELFGHAVSENGA 193
Db 181 NKEKETQVFYSITQGRADTPGVFIERETGMLKYQPLDRQIAKYILFHAVSSNGQ 240
Qy 194 SVEDPMNISIVTDQNDHKPKFTQDTFRGSLVGLPGTSMQVTADEDDAIYTYNGV 253
Db 241 AIEBPEMIVITVDQNDNKPQFTQEVFKASALEGALPGTSMQVTAIDDEVNTYTAAI 300
Qy 254 AYSIHSEPKDHPDLMTIHRSTGTSVISSGDREKVEPYTLTIQATDMDGSGSTTAV 313
Db 301 GYTIPADQDPLPHNKMFTINKETGVISLVTGLDRESFPTTLMVQAADLNGEGLSTAT 360
Qy 314 AVVEIILDANDNAPMFDPOKYEAHPVNAVGHVQRLVTDLDAPNSPAWRATYILIMGGDD 373
Db 361 AVITVLDTNDAFRNPTTYVSGVPENEAIVTTLVTDDADNTPAWEAVTVL-NDN 419
Qy 374 GDHFTITTHPESNOGILTRKGLDFEAKNQHTLYVEVNEAPFVLKPTSTATIVHVED 433
Db 420 EKQFIVVTDVPTNEGTLTKAKGLDFEAKQQYILYVAVTNVAPFEVTLPTSTATVTVDIV 479
Qy 434 VNEAPVFPVPSKVVEVOEGIPTGEPVCVYTAEDPK-ENQKISYRIILRDPAWLMADPDS 492
Db 480 VNEAPIFVPQKRVPEPDEFGVGLSITSYTAREPDTFMEQKITVRIWRDTANWLEINP 539
Qy 493 GOVTAVGTLDRDEBEQFVRNNIYEVNVLAMONGSPPTTGTGTLTLTLIDVNDHGPVPEPRQ 552
Db 540 GAISTRAELDRDVEDVHVKNSTYALIIATDNGSPATGTGTLTLFLDNDVNGVPEPRPT 599
Qy 553 ITICNQSPVRHVLNITDKLSPHTSPPQAQLTDDSDIYWTAEVN-EGDVTVLSKKFLK 611
Db 600 MDFCORNPEPHIININPDLPNTPSPETAELTHGASVNWITIEYNDQRESLILKPKKTLE 659
Qy 612 QDITVDVHLSLSDHGNKEQLTVIRATVCDCHGVETC--PGPWKGGF---LPLVGLAVLA 665
Db 660 LGDHKINKLIDNQNKQDVTTLDVHVCDCDGI VSNCRKARPAPAEAGLQVPAILGLIGLILA 719
Qy 666 LLFLLVLLVLLVLRKKRIKEPLLLPEDDTRDNVYFYEGEGGEDQDYDITQLHRGLEAR 725
Db 720 FLILILVLLVLRVRRVVKPELLPEDDTRDNVYFYDEEGGEDQDQLSGLHGLDAR 779
Qy 726 PEVLRNDVAPTIIPTPMYRPRRPNPDBIGNFIENLKAANTDPTAPPYDILLVFDVEGS 785
Db 780 PEVT-RNDVAPTLMSVQYRPRPNPDBIGNFIDENLKAADSDPTAPPYDLSLLVFDVEGS 838
Qy 786 GSDAASLSSTSSASDQDDYDYLNEGWSRPFKLADMYGGGEDD 829
Db 839 GSEAAATLSLNSSESDDQDDYDYLNEGWRPFKLADMYGGGEDD 882
```



```

RL Proc. Natl. Acad. Sci. U.S.A. 88:11495-11499(1991).
RN [5]
RP DEVELOPMENTAL STAGE.
RC STRAIN=CS7BL/6; TISSUE=Testis;
RA MEDLINE=97033837; PubMed=8879495;
RX Munro S.B., Blaschuk O.W.;
RY "A comprehensive survey of the cadherins expressed in the testes of
RZ fetal, immature, and adult mice utilizing the polymerase chain
RT reaction.";
RL Biol. Reprod. 55:822-827(1996).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 157-370.
RX MEDLINE=96176249; PubMed=8599933; DOI=10.1038/380360a0;
RA Nagar B., Overduin M., Ikura M., Rini J.M.;
RY "Structural basis of calcium-induced E-cadherin rigidification and
RT dimerization.";
RL Nature 380:360-364(1996).
RN [7]
RP STRUCTURE BY NMR OF 157-260.
RX MEDLINE=96271285; PubMed=8785495;
RA Overduin M., Tong K.I., Kay C.M., Ikura M.;
RY "1H, 15N and 13C resonance assignments and monomeric structure of the
RT amino-terminal extracellular domain of epithelial cadherin.";
RL J. Biomol. NMR 7:173-189(1996).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 577-728 IN COMPLEX WITH
RX CTNNB1, AND PHOSPHORYLATION.
RA MEDLINE=21246507; PubMed=11348595; DOI=10.1016/S0092-8674(01)00330-0;
RY Huber A.H., Weis W.I.;
RT "The structure of the beta-catenin/E-cadherin complex and the
RL molecular basis of diverse ligand recognition by beta-catenin.";
CC Cell 105:391-402(2001).
CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types. E-cadherin is a ligand for
CC integrin alpha-E/beta-7.
CC -1- SUBUNIT: Homodimer. Binds CTNNB1.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Non-neuronal epithelial tissues.
CC -1- DEVELOPMENTAL STAGE: In the testis, expression is highest in fetal
CC gonad, then decreases 5-fold in newborn. Detectable in 7-day-old
CC but not in 21-day-old or adult.
CC -1- SIMILARITY: Contains 5 cadherin domains.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X06115; CAA29488.1; -; mRNA.
DR EMBL; X06961; CAA43292.1; -; Genomic DNA.
DR EMBL; X06962; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06963; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06964; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06965; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06966; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06967; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06968; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06969; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06970; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06971; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06972; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06973; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06974; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06975; CAA43292.1; JOINED; Genomic DNA.
DR EMBL; X06339; CAA29645.1; -; mRNA.
DR EMBL; M81449; AAA37352.1; -; Genomic DNA.
DR PIR; S04528; IJMSCE.
DR PIR; S34438; S34438.
DR PDB; 1EDH; X-ray; A/B=156-380.
DR PDB; 1FF5; X-ray; A/B=157-374.
DR PDB; 1I7W; X-ray; B/D=734-884.
DR PDB; 1I7X; X-ray; B/D=734-884.
DR PDB; 1Q1P; X-ray; A=158-369.
DR PDB; 1SUH; NMR; @=156-300.
DR Ensembl; ENSMUSG000000000303; Mus musculus.
DR MGI; MGI:88354; Cdh1.
DR GO; GO:0045177; C:apical part of cell; IDA.
DR GO; GO:0005913; C:cell-cell adherens junction; IDA.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0008013; F:beta-catenin binding; IPI.
DR GO; GO:0005509; F:calcium ion binding; IDA.
DR GO; GO:0016337; P:cell-cell adhesion; IMP.
DR GO; GO:0051260; P:protein homooligomerization; IDA.
DR GO; GO:0019338; P:protein metabolism; IDA.
DR GO; GO:0043026; P:regulation of caspase activation; IDA.
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF000233; Cadherin_C_term.
DR Pfam; PF00028; Cadherin; 5.
DR Pfam; PF01049; Cadherin_C; 1.
DR PRINTS; PR00205; CADHERIN.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS0268; CADHERIN_2; 4.
DR 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
KW Glycoprotein; Phosphorylation; Repeat; Signal; Transmembrane.
FT SIGNAL 1 23 Potential.
FT PROPEP 24 156 Potential.
FT CHAIN 157 884 Epithelial-cadherin.
FT TOPO_DOM 157 709 Extracellular (Potential).
FT TRANSMEM 710 733 Potential.
FT TOPO_DOM 734 884 Cytoplasmic (Potential).
FT DOMAIN 157 264 Cadherin 1.
FT DOMAIN 265 377 Cadherin 2.
FT DOMAIN 378 488 Cadherin 3.
FT DOMAIN 489 595 Cadherin 4.
FT DOMAIN 596 699 Cadherin 5.
FT COMBIAS 840 855 Ser-rich.
FT MOD_RES 840 840 Phosphoserine.
FT MOD_RES 842 842 Phosphoserine.
FT MOD_RES 848 848 Phosphoserine.
FT CARBOHYD 560 560 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 639 639 N-linked (GlcNAc...) (Potential).
FT CONFLICT 267 267 E -> P (in Ref. 2).
FT CONFLICT 272 272 S -> F (in Ref. 2).
FT STRAND 163 166
FT TURN 167 168
FT STRAND 175 179
FT HELIX 183 186
FT TURN 187 187
FT STRAND 190 195
FT STRAND 197 197
FT TURN 198 200
FT STRAND 201 201
FT TURN 204 205
FT STRAND 207 209
FT TURN 211 213
FT STRAND 215 218
FT TURN 224 226
FT STRAND 229 238
FT TURN 239 240
FT STRAND 243 243
FT STRAND 248 255
FT STRAND 263 264
FT TURN 275 276
FT TURN 279 280
FT STRAND 282 285
FT STRAND 288 289
FT TURN 294 296
FT TURN 299 300
FT STRAND 303 310
FT STRAND 319 321

```


Db 249 EIVITVTDQNDNRPEFTQVEFGSVARGAVPGTGVSMKVKSATDADDVVNTYNAIAIYTVS 308

Qy 260 QEPKDPDLFTTHRSVTGTSVSSGLDREKVPETLTIOATMDGSGSTTTAVAVVAIL 319

Db 309 QDPFLPHKNFTVNRDVGTVSVTSGLDRESYPTTLVUQAADLQGBGLSTTAKAVITVK 368

Qy 320 DANDNAPMFQPKYEAEVHPVNAVGVQRIUVTDLDPAPNSPAWRATVLYLMDGGDGDHFTI 379

Db 369 DINDNAPVFNSTYQGVPEVNEVNRATLKVTDGDDAPNTPAWKAVTVV-NDPDQQFVV 427

Qy 380 TTTPESNOGLITRKGLDFAKQHTLYVEVNEAPVVKLPSTATVIVHVDVNEAPV 439

Db 428 VTDPTTNDGLIKTAKGLDFAKQYILHRVENEPEPESGLVPSTATVTVVVDVNEAPI 487

Qy 440 FVPSKVEVQEGIPTCEPVVTAEDPK-ENOKISYRILRDPAGLWANDPDSQGVAV 498

Db 488 FMPAERREVEPDPFGVQGEITSYTAREPDTFMDQKITRYIRWDTANWLEINPETGAIFTR 547

Qy 499 GTLDREDEQFVRNIIYEMVVLAMDNGSPPTTGCTGLTLLTLDVNDHGVPEPRQITICNQ 558

Db 548 AEMDREDAEHKSVTYVALLIATDDGSPATGTGTLVLLVLDVNDNAPIPEPRNMQFCOR 607

Qy 559 SPVRHLVNIITKXLSPTSFPQALTDSDIYMTAEVNERG-DTVVLSLKFKLQDTPYDV 617

Db 608 NPQPHIITILDPDLPNTSPFETALTHGASVNMWTEYNDAAQESLILQPRKDLGEIKYI 667

Qy 618 HLSLDHGNKEQLTVIRATVCDCHGHVETCPGPKGGF-----ILPVLGAVLALLF 668

Db 668 HLKLDNQNDQVTVLTVHVCDCGEGTVNVC---MKAGIVAAAGLVQVPAILGILGAILLI 724

Qy 669 LLVLLVLLVKKRRIKEPLLLPEDDTRDNYVYEGEGGEDDYDITQLHRLGLARPEV 728

Db 725 LILLVLLFLRRVTVKEPLPPDDTRDNYVYDEEGGEDQDFLSQLHRLGLARPEV 784

Qy 729 VLNRDVAPTIPTPMYPRPANPEIGNFIIENKAANTDPTAPPYDTLLVDFYEGSGSD 788

Db 785 T-RNDVAPTLMSVQYPRPANPEIGNFIDENLKAADSDPTAPPYDLSLLVDFYEGSGSE 843

Qy 789 AASLSLTSASDQDDYDLNWSRFRKLLADMYGGGEDD 829

Db 844 AASLSLSLSSDQDDYDLNWSRFRKLLADMYGGGEDD 884

RESULT 12

CADH1 RAT

ID CADH1 RAT STANDARD; PRT; 886 AA.

AC Q9ROT4; Q35794; Q9JIV9;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Epithelial-cadherin precursor (E-cadherin) (Uvomorulin) (Cadherin-1).

CN Name=Cdhl;

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

[1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Sprague-Dawley; TISSUE=Jejunum;

RA Asai K., Tada T., Yamamoto M., Obayashi M., Mizuno M., Toda A.,

RA Eimoto T., Kato T.,

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

[2]

RP NUCLEOTIDE SEQUENCE OF 769-872.

RC STRAIN=Fischer 344; TISSUE=Testis;

RX MEDLINE=20114371; PubMed=10650949; DOI=10.1210/en.141.2.675;

RA Johnson K.J., Patel S.R., Boekelheide K.

RT "Multiple cadherin superfamily members with unique expression profiles

are produced in rat testis.;"

RL Endocrinology 141:675-683(2000).

[3]

RP NUCLEOTIDE SEQUENCE OF 779-846.

RC TISSUE=Mammary tumor;

RA Gibbons K.L.;

RL Thesis (1997), University of Technology / Sydney, Australia.

CC -I- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.

CC They preferentially interact with themselves in a homophilic

CC manner in connecting cells; cadherins may thus contribute to the

CC sorting of heterogeneous cell types. E-cadherin has a potent

CC invasive suppressor role.

CC -I- SUBUNIT: Homodimer. Binds CTNBN1 (By similarity).

CC -I- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).

CC -I- SIMILARITY: Contains 5 cadherin domains.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC -----

DR EMBL; AB017696; BAA84920.1; -; mRNA.

DR EMBL; AF177680; AAF87055.1; -; mRNA.

DR EMBL; AJ000540; CAA04173.1; -; mRNA.

DR HSP; P09803; I17X.

DR SMR; Q9ROT4; 159-376, 806-881.

DR Ensembl; ENSRNOG0000020151; Rattus norvegicus.

DR GO; GO:0007416; P:synaptogenesis; IMP.

DR InterPro; IPR002126; Cadherin.

DR InterPro; IPR000233; Cadherin_C_term.

DR Pfam; PF00028; Cadherin; 5.

DR Pfam; PF01049; Cadherin_C; 1.

DR PRINTS; PR0205; CADHERIN.

DR SMART; SM00112; CA; 4.

DR PROSITE; PS00232; CADHERIN_1; 3.

DR PROSITE; PS0268; CADHERIN_2; 5.

KW Calcium; Cell adhesion; Glycoprotein; Phosphorylation; Repeat; Signal;

KW Transmembrane.

FT SIGNAL 1 23 Potential.

FT PROPEP 24 158 Potential.

FT CHAIN 159 886 Epithelial-cadherin.

FT TOPO_DOM 24 713 Extracellular (Potential).

FT TRANSHEM 714 734 Potential.

FT TOPO_DOM 735 886 Cytoplasmic (Potential).

FT DOMAIN 159 266 Cadherin 1.

FT DOMAIN 267 379 Cadherin 2.

FT DOMAIN 380 490 Cadherin 3.

FT DOMAIN 491 597 Cadherin 4.

FT DOMAIN 598 701 Cadherin 5.

FT COMPTAS 842 857 Ser-rich.

FT MOD_RES 842 842 Phosphoserine (By similarity).

FT MOD_RES 844 844 Phosphoserine (By similarity).

FT MOD_RES 850 850 Phosphoserine (By similarity).

FT CARBOHYD 562 562 N-linked (GLCNAC. .) (Potential).

FT CARBOHYD 641 641 N-linked (GLCNAC. .) (Potential).

SQ SEQUENCE 886 AA; 98715 MW; A9AE28EB797A547 CRC64;

Query Match 54.7%; Score 2388; DB 1; Length 886;

Best Local Similarity 53.8%; Pred. No. 1.9e-140;

Matches 477; Conservative 122; Mismatches 212; Indels 76; Gaps 14;

Qy 9 ASLLLLQV-CWLQC-----AASEPQAVFRAEAVTELAGGAQEPQALGKV-FMGCPQGE 62

Db 10 ALLLLLQVSSWL-CQPESESDSCRPGFSSEVYTLVPERHLERGHILGRVKFEGCTGRP 68

Qy 63 PALFSTNDPDTVRNGETVQERRSLK----- 88

Db 69 RTAFSEDSRPFKVSTGDGVITVKRLKHLKHLKLETSFLVHAWDSSYRKLSTKVTLKSLGHHHH 128

Qy 89 -----ERNP-LKIFFPS-KRILRRHKRDVWVAPISVPENKGKPPQRLNQKSNKDR 137

Db 129 RHHRDVPVSENPELLTFPSFHQGLRQKRDWVIPPINCENQKGEPPQRLVQIKSNRDK 188

Qy 138 DTKIFYSITGPGADSPGEGVFAVEKETGWLILNKPLDREETAKYELFGHVAENGASVED 197

Db 189 ETVFYISITGPGADKPPGVGVFIIBRETGWLKTQFLDREAIDKYLILYSHAVSNGEAVED 248

QY 198 PMNISIIVTDONHKKFTQDTRGSLVGLPCTSMQVATATDEDDAIYTYNGVAVYSI 257
Db 249 PMEIVVTVDONDRBEFIEQVEFGSVAEGALPCTSMQVATADDDINTYNAAYTI 308
QY 258 HSQEPKDPHLMFTIHRSTGTSIVISGLDREKVPYTLTIQATDMDGSGSTTAVAVVE 317
Db 309 LSQDPELPHKMTFVARDTGIVSVTSGLDRESYPTTLLVVQADLQGBGLSTTAKAVIT 368
QY 318 ILDANDNAPMDFQKVEAHVPENAVGHEVORLVTDLDPNSPAWRATYLMIGDGDGH- 376
Db 369 VKDINDNAPINFSTVQGVLENEVGARIATLKVTDDDAPNTAWNAVTVV--NPDDHQ 426
QY 377 FTITTHPESNQGLITTRKGLDFAKQHTLYVEVNEAPEVLKLPSTSTATIVVHVVDVNE 436
Db 427 FTVITDPKTNIGLTKAGLDLFAKQYIHLHVNEEPEPESLVPSTATVTVVDVVDVNE 486
QY 437 APVFPVPSKVVEQGIPTGPEVCVVTAEPPDK-ENQKISYRILRDPAGHLANDPDSGV 495
Db 487 APIFVPAEKRVPEPFGVLEIASYTAREPDTFMEQKITRYIRWDTANWLEINPETGVI 546
QY 496 TAVGTLDREDEQVRNNIYEVMLAMNDGSPPTTGTGTLTLLTLDVNDHGPVPEPRQITI 555
Db 547 STRAENDREDSERHKSTYATLIIATDGCSPATGTGTGTLTLLVSDVNDNAPPEPRNQF 606
QY 556 CNOQSPVRHVNITDKLSPHSTSPFQALTDSDSDIYWTAEVNE-EGDTVVLKXKFLKQDT 614
Db 607 CQNPXPHVITILDPOLPNTSPFATLTHGASVNWNTIEYNDAEQBSLILQPKDKLEIGS 666
QY 615 YDVHLSLSHGNKEQLTVIRATVCDCHGHVETC-----PGPKKGGFILPVGA 662
Db 667 YKINKLSLONQKQDVTLEHVHVCDEGTVNNCMKAISLQVPA-----ILGILGG 720
QY 663 VLALLFLLVLLLVKRRKIKIPELLLPEDDTRDNNVYFYEEGEGGEDDQDITQLHRGL 722
Db 721 ILALLLILLLLVFLRRRTVWKEPLFPDDTRDNNVYFYEEGEGGEDDQDPLSQLHRGL 780
QY 723 EAPPEVLNDVAPTIIPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDYLAVFY 782
Db 781 DARPEVI-RNDVAPTLMSVQYRPRPANPDEIGNFIENLKAANDSPTAPPYDYLAVFY 839
QY 783 EGSQSDAASLSSTSSASODQDQDYLVNWSRPFKGLADMYGGGEDD 829
Db 840 EGSQSEAASSLSSNESDQDQDYLVNWSRPFKGLADMYGGGEDD 886

RESULT 13
Q5RAX1_PONPY
ID Q5RAX1_PONPY PRELIMINARY; PRT; 882 AA.
AC Q5RAX1;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKF7p469A211.
GN Name=DKF7p469A211;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=kidney;
RG The German cDNA Consortium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Oeinger A., Fobg G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR858890; CAH91089.1; -; mRNA.
DR SNR; Q5RAX1; 155-371.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007156; P:phospholipid cell adhesion; IEA.
DR InterPro; IPR002126; Cadherin.

DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; Cadherin; 5.
DR Pfam; PF01049; Cadherin C; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00122; CA; 4.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS0268; CADHERIN_2; 4.
KW Calcium; Hypothetical protein.
SQ SEQUENCE 882 AA; 97427 MW; 3EC35657C1B6C719 CRC64;

Query Match 54.6%; Score 2387; DB 2; Length 882;
Best Local Similarity 54.5%; Pred. No. 2.2e-140; Indels 60; Gaps 13;
Matches 478; Conservative 114; Mismatches 225;

QY 9 ASILLIQV-CWLQCAASEPCRAVFREAETVLEAGAEQEPGQALGV-FWGCPG-QEPAL 65
Db 10 ALLLLLQVSSWL-CQEPFCHGFDASVYTFVPRHLGRVILGRVFNEDCTGRQRTAY 68
QY 66 FSTDNDFTVRNGETVQERRSLKERNP----- 92
Db 69 FSLDT-RFKVGTGDTVTKRPLRFHNPQIHFLVYAWDSTYRKPFSTKVTLTNTVGHHRPLP 127
QY 93 -----LKIIP-SKRILRRHKRDWVAPISVPENKGGPFQRLNOLKSNKDRDTK 140
Db 128 HQASVSGIQAEALTTFNSSGSLRRRRKRDWVIPPISCPENEKGGPFQRLNOLKSNKDRDTK 187
QY 141 IFYSITGPGADSPPGCVFAVEKETGMLLNLKPLDREIAKYELFHAVSSNGASVEDPMN 200
Db 188 VFYSITGPGADTPVGVFIERETGWLKVTPELDRERIATYTLFHAVSSNGASVEDPME 247
QY 201 ISIVTDQNDHKPKFTQDTPFRGSLVGLVPGTSMQVATATDEDDAIYTYNGVAVYSIQ 260
Db 248 ILITVTDQNDKPEFTQEVFKGSMEGALPGTSVMEVATADDDVNTYNAAYTILSQ 307
QY 261 EPKDPHLMFTIHRSTGTSIVISGLDREKVPYTLTIQATDMDGSGSTTAVAVVEILD 320
Db 308 DPDELQKNNFTINRNTGVISVVTGLDRESFTYTLVQQAADLQSGELSTTATVITVD 367
QY 321 ANDNAPMDFQKVEAHVPENAVGHEVORLVTDLDPNSPAWRATYLMIGDGDGHFTIT 380
Db 368 TNDNPFVNPTTYKQVPEDEANVTTLKVTDADAPSTPAWEAVYTIL-NDNGQGFVVT 426
QY 381 THPESNQGLITTRKGLDFAKQHTLYVEVNEAPEVLKLPSTSTATIVVHVVDVNEAPVF 440
Db 427 TNPVNDGILTKAGLDLFAKQYIHLVAVTNVVPFVSLTSTATVTVVDVLDVNEAPIF 486
QY 441 VPSPKVEVQEGIPTGPEVCVVTAEPPDK-ENQKISYRILRDPAGHLANDPDSGVTA 499
Db 487 VPPEKRVESDFGVGQEIYTSYTAWEPTFMEQKITRYIRWDTANWLEINPDTGAISTRA 546
QY 500 TLDREDEQVRNNIYEVMLAMNDGSPPTTGTGTLTLLTLDVNDHGPVPEPRQITICNS 559
Db 547 ELDRDVEHVKNSTYATLIIATDNGSPVATGTGTLTLLVSDVNDNAPPEPRTLFCERN 606
QY 560 PVRHVNITDKLSPHSTSPFQALTDSDSDIYWTAEVNE-EGDTVVLKXKFLKQDTYDVH 618
Db 607 PKPQVINIADLPNTSPFATLTHGASVNWNTIEYNDPTQBSIILKPKMALEVGVDYKIN 666
QY 619 LSLSDHGNKEQLTVIRATVCDCHGHVETC--PGPKKGGF-----ILPVGLAVALLLVLV 672
Db 667 LKLMONQNKQDVTLEVGVCDCGCVAGVCKKQAPTEAGLQIPAILGILGILLALLLIL 726
QY 673 LLLLVKRRKIKIPELLLPEDDTRDNNVYFYEEGEGGEDDQDITQLHRGLRARPVVLNRN 732
Db 727 LLLFLRRRAVKEPLLPEDDTRDNNVYFYEEGEGGEDDQDPLSQLHRGLDARPEVT-RN 785
QY 733 DVAPTIPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDYLAVFYEGSGSDAASL 792
Db 786 DVAPTLMSVPRYLPRPANPVEIGNFIENLKAANDTPTAPPYDYLAVFYEGSGSEAA 845
QY 793 SSLTSSASODQDQDYLVNWSRPFKGLADMYGGGEDD 829
Db 846 SSLNSESSEKDDQDQDYLVNWSRPFKGLADMYGGGEDD 882

S -> P (in Ref. 2).
S -> P (in Ref. 2).

573	573
864	864
165	162
178	174
183	183
185	189
196	189
199	197
200	200
204	203
208	206
212	210
217	214
225	223
237	228
242	242
254	247
263	262
273	268
275	274
284	280
288	287
299	298
308	302
320	318
324	322
329	326
338	336
349	341
353	350
354	354
359	357
367	362
376	376
386	380
397	393
399	399
406	405
409	408
412	412
413	414
416	416
420	420
421	420
428	424
432	430
439	435
455	451
459	458
479	472
487	486
492	491
496	496
499	498
504	504
511	511
517	516
526	522
537	535
541	542
545	548
552	551
569	562
582	579
602	609
614	610
623	620
629	639
642	642
645	644
652	648
665	662
667	665

FT	STRAND	676	679	
FT	STRAND	682	684	
SQ	SEQUENCE	880 AA; 97651 MW; C6CCD91566427D86 CRC64;		
Query Match 54.4%; Score 2375; DB 1; Length 880;				
Best Local Similarity 54.2%; Pred. No. 1.2e-139;				
Matches 471; Conservative 128; Mismatches 214; Indels 56; Gaps 9;				
Qy	11	LLLLQVCLQCAASEPRAREVLEAGAEQERQALGV-FWGCGPQEPALFSTD	69	
Db	17	LCLLQVPSINADVSGCKPGFSSAEYIFSVNRRELERGRKLGVNFSDCTTRKHGLYVG	76	
Qy	70	NDDFTV-----RNGETVQERRS	86	
Db	77	DSRFRVLPGTVLVKRVHKLKDKFTISTWDARGIKHSTNIAVASKRHRSGEEAHSRS	136	
Qy	87	LKERNPLKIPPSKRI-LRRHKRWVWAPIVSPENGKGPFFORLNQLKSNKORDTKIFYSI	145	
Db	137	---KLPVLTFPETHTGLKRRKRDWVIPPVKYSENERGPFPRKLVQIKSNKDRFNKVYYSI	193	
Qy	146	TGPCADSPPEGVFAVEKETGWLKLPDRREEIAKVELFGHAVSENGASVEDPMNISIIV	205	
Db	194	TQGADAPPQGVFRIEWETGMLVTRPDRREYDKYVLSHAVSENGSPVEEPMETINV	253	
Qy	206	TDQNDHKPKFTQDTRFGSVLEGVLPGTSVMQVATDEDDAIYTYNGVVAYSIHSQEPKOP	265	
Db	254	IDQNDNRPKFTQDVFRGSVREGVQGTQVMASATDEDDNIDSLNGLSVSILKQDPEEP	313	
Qy	266	HLMFTIHRSTGTTISVSSGLDRKVPYTLTIQATMDGSGSTTTAVAVVEILDANDNA	325	
Db	314	IPNLFTINRETGVISLIGTGLDRKFPYTLTVQATDLEGAGLSVEGKAIQITDANDNA	373	
Qy	326	PMFDPQKVEAHVPENAVCHVEORLTVDLDAPNSPANRATYILMGDDGDGHFTITTHPES	385	
Db	374	PIFDPKTYTALVPENEIGFEVQRLSVTDLOMPGTAWQAVYKIR-VNEGGFNITTDPE	432	
Qy	386	NOGILTRKGLDFAKNOHTLYEVETNEAPPVLKLPSTATIVVHVVEDVNEAPVFVPPSK	445	
Db	433	NOGILTTAKGLDFELRKQVLYQITVENAEPSPVPLPTSTATVTVVEDVNEAPFFVPAVS	492	
Qy	446	VVEVQGIPTGEPVCVYTAEDPKEN-QKISYRILRDPAGWLAMPDSGGVAVGTLDRE	504	
Db	493	RVDVSEDLGRKEIISLVAQDPDKQIQKLSYFICNDPARWLTVNKONGIVTGNGLDRE	552	
Qy	505	DEQFVRNNIYEMVLAMDNGSPPTGTGTLTLTDVNDHGPVPEPRQITICNOSPVRHV	564	
Db	553	SE-YVKNNITYTIVMLVTDGVSVGTGTGLTLHLVLDVNDNGVPSPRVFTMCDQNPEPQV	611	
Qy	565	LNITDKLSPTSPPQALTDSDIYMTAEVNEEGDVTVLSLKKFLKQDTYDVHLSLSDH	624	
Db	612	LTISDADIPPNTYKYVLSHGSDLTWKAELDSKTSMLLSPTQQLKKGYSIYVLSDA	671	
Qy	625	GKKEQLTVIRATVCDCHGHVETCPGPMKGGFILP-----VLGAVLALLFLLLVLLLRKX	680	
Db	672	QNNPQLTVVNATVCSCEGKAIKQEKLVGGFDLPILVILGSLVALLILFLLLLLFLKRX	731	
Qy	681	RKIKEPLLPEDDTRDNVYVYEGEGGEEDDYDITOLHGLEARPEVWLENDVAPTIIP	740	
Db	732	KVWKEPLLPEDDTRDNIFYYEGEGGEEDDYDLSQLHRLGRDP-IMRNDVVPTLMP	790	
Qy	741	TPMYRPRPANDPEIGNFTIENLKAANTDPTAPPYDTLLVFDYEGSGSDAASLSLTSSAS	800	
Db	791	APHYRPRPSNDEIGNFTIDENLDAANDPTAPPYDSLLVFDYEGSGSENASLSLSNSNS	850	
Qy	801	DQDQDYDYNLWNGSRFKKLADMYGGEDD	829	
Db	851	NDEHDYNYLSDWGSFRKRLADMYGGDDDE	879	


```
Db      1  MELLSGPHAFLLLLQVCLWLSVVSSEPYRAGFIGEAGVTLEVEGTDLEPSQVLGKVALAG 60
Qy      59  PGQEPALFSTNDNDDFTVRNGETQOERSLKERNPLKIFPSKRILRRHKRDVWVAPISVPE 118
Db      61  QGMHHA----DNGDIIMLTRGTVOGGDAMHS-----PPTRIILRRKRREWMPPIFVPE 110
Qy     119  NGKGPFFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPDLREEI 178
Db     111  NGKGPFFQRLNQLKSNKDRGTKIFYSITGPGADSPPEGVFTIEKESGWLILHHPDLREKI 170
Qy     179  AKYELFHAVSENGASVEDPMNISIIITDQNDHKPKFTQDTRFGSVLEGLVPGTSVMQVT 238
Db     171  VKYELYGHAVSENGASVEEPMNISIIITDQNDNKPFTQDTRFGSVIEGVMPGTSVMQVT 230
Qy     239  ATDEDDAIYTYNGVAVSIHSEKPKDPLMFTIHRSTGTISVSSGLDREKPEYRLTI 298
Db     231  ATDEDDAVTYNGVAVSIHSEKPKDPLMFTIHRSTGTISVSSGLDREKPEYRLTV 290
Qy     299  QATDMGDSGTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHEVQRLTVTDLDPN 358
Db     291  QATDMGEGSTTTAAVAVVQLDANDNAPEPEPOKYEAVPENEVGEVQRLTVTDLDPN 350
Qy     359  SPAWRATYILMGDDGDHFTITTHPESNQGLITRKGDLFEAKNQHTLYVEVTNEAPFVL 418
Db     351  WPAWRATYHIVGGDDGDHFTITTHPETNQGLITRKGDLFEAQDQHTLYVEVTNEAPFAV 410
Qy     419  KLPTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTEGPEVCVVTADDPKENOKISYRI 478
Db     411  KLPTATATVVHVVKDNEAPVFPVPSKVIEAQEGISIGELVCIVTAQDPKEDOKISYTI 470
Qy     479  LRDPAGWLAMPDPSGOVTAVGTGLDREDEQFVRNNIYEVWVLAMDNGSPPTTGTGLLTL 538
Db     471  SRDPANWLAVDPDSGOITAAGLDREDEQFVKNNVYEVWVLATDSGNPPTTGTGLLTL 530
Qy     539  IDVNDHGFVPEPROITTCNOSPVRHVLNITDKOLSPHTSPFQAQLTDDSDIYWTAEVNEE 598
Db     531  TDINDHGPIPEPRQIICNOSPVPQVLNITDKOLSPNSPFAQLTHDSDIYWMAEVSEK 590
Qy     599  GDTVVLISKKFLKODTYDVHLSLSDHGNKQOLTVIRATVCDCHGV-ETCPGPKGGFIL 657
Db     591  GDTVALSKKFLKODTYDLHLSLSDHGNREQLTMIRATVCDCHGOVFNDCPRPKGGFIL 650
```

RESULT 2

```
US-08-474-068A-9
; Sequence 9, Application US/08474068A
; Patent No. 5837525
; GENERAL INFORMATION:
; APPLICANT: Ranscht, Barbara
; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,068A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/213,361
; FILING DATE: 14-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/607,293
; FILING DATE: 30-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1683
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-474-068A-9

Query Match      81.3%; Score 3553.5; DB 1; Length 822;
Best Local Similarity 81.5%; Pred. No. 0;
Matches 678; Conservative 58; Mismatches 83; Indels 13; Gaps 5;

Qy      1  MGLPRGPLA-SLLLLQVCLWLSVVSSEPYRAGFIGEAGVTLEVEGTDLEPSQVLGKVALAG 60
Db      1  MELLSGPHAFLLLLQVCLWLSVVSSEPYRAGFIGEAGVTLEVEGTDLEPSQVLGKVALAG 60
Qy     59  PGQEPALFSTNDNDDFTVRNGETQOERSLKERNPLKIFPSKRILRRHKRDVWVAPISVPE 118
Db     59  PGQEPALFSTNDNDDFTVRNGETQOERSLKERNPLKIFPSKRILRRHKRDVWVAPISVPE 118
Qy     61  QGMHHA----DNGDIIMLTRGTVOGGDAMHS-----PPTRIILRRKRREWMPPIFVPE 110
Db     61  QGMHHA----DNGDIIMLTRGTVOGGDAMHS-----PPTRIILRRKRREWMPPIFVPE 110
Qy    119  NGKGPFFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPDLREEI 178
Db    119  NGKGPFFQRLNQLKSNKDRDTKIFYSITGPGADSPPEGVFAVEKETGWLILNKPDLREEI 178
Qy    111  NGKGPFFQRLNQLKSNKDRGTKIFYSITGPGADSPPEGVFTIEKESGWLILHHPDLREKI 170
Db    111  NGKGPFFQRLNQLKSNKDRGTKIFYSITGPGADSPPEGVFTIEKESGWLILHHPDLREKI 170
Qy    179  AKYELFHAVSENGASVEDPMNISIIITDQNDHKPKFTQDTRFGSVLEGLVPGTSVMQVT 238
Db    171  VKYELYGHAVSENGASVEEPMNISIIITDQNDNKPFTQDTRFGSVIEGVMPGTSVMQVT 230
Qy    239  ATDEDDAIYTYNGVAVSIHSEKPKDPLMFTIHRSTGTISVSSGLDREKPEYRLTI 298
Db    231  ATDEDDAVTYNGVAVSIHSEKPKDPLMFTIHRSTGTISVSSGLDREKPEYRLTV 290
Qy    299  QATDMGDSGTTTAVAVVEILDANDNAPMFDPOKYEAHVPENAVGHEVQRLTVTDLDPN 358
Db    291  QATDMGEGSTTTAAVAVVQLDANDNAPEPEPOKYEAVPENEVGEVQRLTVTDLDPN 350
Qy    359  SPAWRATYILMGDDGDHFTITTHPESNQGLITRKGDLFEAKNQHTLYVEVTNEAPFVL 418
Db    351  WPAWRATYHIVGGDDGDHFTITTHPETNQGLITRKGDLFEAQDQHTLYVEVTNEAPFAV 410
Qy    419  KLPTSTATIVVHVEDVNEAPVFPVPSKVVEQEGIPTEGPEVCVVTADDPKENOKISYRI 478
Db    411  KLPTATATVVHVVKDNEAPVFPVPSKVIEAQEGISIGELVCIVTAQDPKEDOKISYTI 470
Qy    479  LRDPAGWLAMPDPSGOVTAVGTGLDREDEQFVRNNIYEVWVLAMDNGSPPTTGTGLLTL 538
Db    471  SRDPANWLAVDPDSGOITAAGLDREDEQFVKNNVYEVWVLATDSGNPPTTGTGLLTL 530
Qy    539  IDVNDHGFVPEPROITTCNOSPVRHVLNITDKOLSPHTSPFQAQLTDDSDIYWTAEVNEE 598
Db    531  TDINDHGPIPEPRQIICNOSPVPQVLNITDKOLSPNSPFAQLTHDSDIYWMAEVSEK 590
Qy    599  GDTVVLISKKFLKODTYDVHLSLSDHGNKQOLTVIRATVCDCHGV-ETCPGPKGGFIL 657
Db    591  GDTVALSKKFLKODTYDLHLSLSDHGNREQLTMIRATVCDCHGOVFNDCPRPKGGFIL 650
Qy    658  PVLGAVLALLFLLVLLVLRKKRKEPFLLLPDDTRDNVYFYGEGGGEEDQDYDITQ 717
Db    651  PILGAVLALLTLLALLVLRKKRKEPFLLLPDDTRDNVYFYGEGGGEEDQDYDITQ 710
Qy    718  LHRGLEARPEVLRNDVAPTIPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDTL 777
Db    711  LHRGLEARPEVLRNDVVPTEPTPMYRPRPANPDEIGNFIENLKAANTDPTAPPYDSL 770
Qy    778  LVFDYEGSGDAASLSSLTSSASQDQDYDYLNEGSRFKLADMYGGGEDD 829
Db    771  MVFDYEGSGDAASLSSLTSSASQDQDYDYLNEGSRFKLADMYGGGEDD 822
```


QY 658 PVLGAVLALLFLLVLLVLLVRRKKRKEPILLPDDTRDNVFFYBEGGEGEDDQDYDTQ 717
DB 651 PILGAVLALLTLLALLVLLVRRKKRKEPILLPDDTRDNVFFYBEGGEGEDDQDYDTQ 710
QY 718 LHRGLEARPEVLRNDVAPITIIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDIL 777
DB 711 LHRGLEARPEVLRNDVAPITIIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDSL 770
QY 778 LVFDYEGSGDAASLSLTSASDQDQDYDYLNEWGRFKKLADMYGGGEDD 829
DB 771 MVFDYEGSGDAASLSLTSASDQDQDYDYLNEWGRFKKLADMYGGGEDD 822

RESULT 3
US-08-472-481-8
; Sequence 8, Application US/08472481
; Patent No. 5863804
; GENERAL INFORMATION:
; APPLICANT: Ranscht, Barbara
; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,481
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/213,361
; FILING DATE: 14-MAY-1994
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1696
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-472-481-8

Query Match 81.3%; Score 3553.5; DB 1; Length 822;
Best Local Similarity 81.5%; Pred. No. 0;
Matches 678; Conservative 58; Mismatches 83; Indels 13; Gaps 5;

QY 1 MGLPRGPLA-SLLLLQVWLQCAASEPCRAVVF-REAEVTLAAGAEQPGQALGKVFMG 58
DB 1 MELLSGPHAFLLLLQVWRSVSEPYRAGFTGEAGVTLLEVGTDLPSQVLGKVALAG 60
QY 59 PGQEPALFSTDNDQDFTVRNGETQERRSLKERNPLKIFPSKRILRRHRKRDVWVAPISVPE 118
DB 61 QGMHHA---DNGDIIMLTGRGTQGGKAMHS-----PPTRILRRKRREWVMPPIFVPE 110
QY 119 NGKGPPPPQRLNQLKSNKDRQTKIFYSITGPGADSPPEGVFAVEKETGWLILLNPLDREI 178
DB 111 NGKGPPPPQRLNQLKSNKDRQTKIFYSITGPGADSPPEGVFTIEKESGWLILLHMLDREKI 170

QY 179 AKYELFGHAVSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGSLVLEGLVPGTSMQVT 238
DB 171 VKYELVGHAVSENGASVEPMNISIIITDQNDNKPKFTQDTFRGSLVIEGMPGTSVMQVT 230
QY 239 ATDEDDAIYTYNGVAYSISHQEPKDPHDMETIHRSTGCTISVISSGLDREKVPETLII 298
DB 231 ATDEDDAVNTYNGVAYSISHQEPKDPHDMETIHKSTGCTISVISSGLDREKVPETLII 290
QY 299 QATDMGDGSTTTAVAVVEILDANDNAPMFDQPKYEAHVPENAVGEVORLTVTDLDAFN 358
DB 291 QATDMGEGSTTTAEAVVOILDANDNAPFEFEQKYEAVPEVGEVORLTVTDLDPVN 350
QY 359 SPARATYILIMGGDDGDHFTIITHPESNQILITTRKGLDPEAKNQHTLVYVETNEAPFVL 418
DB 351 WPAWRATYHIVGGDDGDHFTIITHPETNQVLTTKKGLDPEAQDQHTLVYVETNEAPFAV 410
QY 419 KLPTSTATTIVHVEDVNEAPVFPVPSKVVEGEGIPTGEPCVYTAEDPOKNOKISYRI 478
DB 411 KLPTATATVVHVHVDVNEAPVFPVPSKVIEAQEGISIGELVCITYTAQDPKEDQKISYTI 470
QY 479 LRDPAGWLAMPDPSGOVTAVGTLDREDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLL 538
DB 471 SRDPANWLAVDPSGOITRAGILDREDEQFVKNVYEVNVLATDSGNPPTTGTGTLTLL 530
QY 539 IDVNDHGPVPEPRQITICNQSPVRHVLNITDKDLSPTSPFQAQLTDDSDIYVTAENEE 598
DB 531 TDINDHGPIPEPRQIIICNQSPVQVLNITDKDLSNPSPPFQAQLTHDSDIYMAESVSEK 590
QY 599 GDTVLSLKKFLKQDTYDVHLSLSDHGNKEQLTVIRATVCDCHGV-ETCPGPKGGFIL 657
DB 591 GDTVALSLKKFLKQDTYDLHLSLSDHGNREQLTMRATVCDCHGVFNDCPRPWKGGFIL 650
QY 658 PVLGAVLALLFLLVLLVLLVRRKKRKEPILLPDDTRDNVFFYBEGGEGEDDQDYDTQ 717
DB 651 PILGAVLALLTLLALLVLLVRRKKRKEPILLPDDTRDNVFFYBEGGEGEDDQDYDTQ 710
QY 718 LHRGLEARPEVLRNDVAPITIIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDIL 777
DB 711 LHRGLEARPEVLRNDVAPITIIPTMYRPRPANPDEIGNFIENLKAANTDPTAPPYDSL 770
QY 778 LVFDYEGSGDAASLSLTSASDQDQDYDYLNEWGRFKKLADMYGGGEDD 829
DB 771 MVFDYEGSGDAASLSLTSASDQDQDYDYLNEWGRFKKLADMYGGGEDD 822

RESULT 4
US-08-474-067-8
; Sequence 8, Application US/08474067
; Patent No. 5811518
; GENERAL INFORMATION:
; APPLICANT: Ranscht, Barbara
; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,067
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/213,361
; FILING DATE: 14-MAY-1994


```
QY 200 NISIIVTQNDHKKPKTQDTFRGSLVLEGLVPGTSVMQVATDEDDDAIYTYNGVAVYSIHS 259
Db 249 EIVITVTDQNDNRPEFTQEVFEGSVAEGAVPGTSVMKVSATDADDVNTYAAIAYTIVS 308
QY 260 QEPKDPHDLMTFTHRTGTISVSSGLDRKVEYTLTIQATDMGDGTTTAVAVETL 319
Db 309 QDPFLPHKMFVNRTGTIVSVTSGLDRESYTYTLVQAAADLQEGSLTAKAVITVK 368
QY 320 DANDNAPMFDPOKYEAHVPENAVGVHEVQRLVTVDLDAPNSPAWRATYILMGDDGDGHFTI 379
Db 369 DINDNAPVFNSTYQGVPENEVNARIATLKVTDADAPNTPAWKVYTVV-NDPQOQFVV 427
QY 380 TTHPESNOGILTRKGLDPEAKNQHTLYVEVTNEAPFVLKLTSTATIVVHVVEDVNEAPV 439
Db 428 VTDPTTNDGILKTAKGLDPEAKQVYLHVRENEEPFEGSLVPSTATVTVVVDVNEAPI 487
QY 440 FVPPSKVVEVQEGIPTEGVCVYTAEDDPK-ENOKISYRILRDPAGWLAMDPSGOVTAV 498
Db 488 FMPAERRVEVPDFGVQGEITSYTAREPDTFMDQKITRYIWRDTANWLSEINPETGAIFTR 547
QY 499 GTLDREDOFVRNNIYEVNVLAMDNGSPPTGTGTLTLLTLDVNDHGPVPEPRQITICNQ 558
Db 548 AEMDREDAEHVKNSTYVALIATDDGSPATGTGTLTLLVLDVNDNAPPEPRNQFCQR 607
QY 559 SPVRHVLNITDKLSFHTSPFOALTDSDSIYWTAEVNEEG-DTVVLSKKFLKQDTYDV 617
Db 608 NPQPHIITILDPDLPNTSPFTAELTHGASVNWNTIEYNDAQESLILQPRKOLEIGEYKI 667
QY 618 HLSLDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGF-----ILPVLGAVLALLF 668
Db 668 HLKLDNQNDQVTTLDVHVCDCEGTNNC---MKAGIVAAGLQVPAIILGILGILLALI 724
QY 669 LLLVLLLVKRRKIKPELILPEDDTRDNVYVYEGEGGEEDQDYITOLHRLGLEARPEV 728
Db 725 LILLLLFLLRRRTVVEKPLPPDDTDRDNVYVYDEEGGEEDQDFLSQLHRLGLDARPEV 784
QY 729 VLRNDVAPTIPTPMYRPRPANPEIGNFTIENLKAANTDPTAPPYDTLLVDFYEGSGSD 788
Db 785 T-RNDVAPTLMSVQYRPRPANPEIGNFTIENLKAADSDPTAPPYDLSLLVDFYEGSGSE 843
QY 789 AASLSLTSASDQDYDYLNEWGRPKLADMYGGEDD 829
Db 844 AASLSLSNSESQDQDYDYLNEWGRPKLADMYGGEDD 884
```

RESULT 6

```
US-08-472-481-7
; Sequence 7, Application US/08472481
; Patent No. 5863804
; GENERAL INFORMATION:
; APPLICANT: Ranecht, Barbara
; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472, 481
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/213,361
; FILING DATE: 14-MAY-1994
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 07/607,293
; FILING DATE: 30-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1686
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 884 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-472-481-7

Query Match 54.6%; Score 2384; DB 1; Length 884;
Best Local Similarity 53.5%; Pred. No. 4,1e-213;
Matches 471; Conservative 123; Mismatches 221; Indels 66; Gaps 13;

QY 9 ASLLLLQV-CWLOCAASEP--CRAVFREAETLEAGGAEQEPQALGVK-FMGCPEQPEA 64
Db 10 ALLLLQLQVSSWL-CQELEPESCSPGFSSEVYTFPVERHLERGHVLRVRFEGCTGRPT 68
QY 65 LFSTNDPFTVRNGTVOERRSLK----- 88
Db 69 AFFESDRSPKATDGTITVKRHLKLKLETSLFVRAROSSHRELSTKVTLKSMGHHHRH 128
QY 89 -----RNP-LKIFPSPKRI-LRRHKRDWVAPISVPENKGGPPORLNQKSNKDRDT 139
Db 129 HHRDPASENPELLMFPSPVGLRRQKRDWVIPPISCPENKGEFPKVLQIKSNRKET 188
QY 140 KIFYSITGPGADSPPEGVFAVEKETGWLILNKPLDREETIAKYLEFHGAVSENGASVEDPM 199
Db 189 KIFYSITGQADKPPGVFIIERETGWLKVQPLDREALAKYILSHAVSSNGEAVEDEM 248
QY 200 NISIIVTQNDHKKPKTQDTFRGSLVLEGLVPGTSVMQVATDEDDDAIYTYNGVAVYSIHS 259
Db 249 EIVITVTDQNDNRPEFTQEVFEGSVAEGAVPGTSVMKVSATDADDVNTYAAIAYTIVS 308
QY 260 QEPKDPHDLMTFTHRTGTISVSSGLDRKVEYTLTIQATDMGDGTTTAVAVETL 319
Db 309 QDPFLPHKMFVNRTGTIVSVTSGLDRESYTYTLVQAAADLQEGSLTAKAVITVK 368
QY 320 DANDNAPMFDPOKYEAHVPENAVGVHEVQRLVTVDLDAPNSPAWRATYILMGDDGDGHFTI 379
Db 369 DINDNAPVFNSTYQGVPENEVNARIATLKVTDADAPNTPAWKVYTVV-NDPQOQFVV 427
QY 380 TTHPESNOGILTRKGLDPEAKNQHTLYVEVTNEAPFVLKLTSTATIVVHVVEDVNEAPV 439
Db 428 VTDPTTNDGILKTAKGLDPEAKQVYLHVRENEEPFEGSLVPSTATVTVVVDVNEAPI 487
QY 440 FVPPSKVVEVQEGIPTEGVCVYTAEDDPK-ENOKISYRILRDPAGWLAMDPSGOVTAV 498
Db 488 FMPAERRVEVPDFGVQGEITSYTAREPDTFMDQKITRYIWRDTANWLSEINPETGAIFTR 547
QY 499 GTLDREDOFVRNNIYEVNVLAMDNGSPPTGTGTLTLLTLDVNDHGPVPEPRQITICNQ 558
Db 548 AEMDREDAEHVKNSTYVALIATDDGSPATGTGTLTLLVLDVNDNAPPEPRNQFCQR 607
QY 559 SPVRHVLNITDKLSFHTSPFOALTDSDSIYWTAEVNEEG-DTVVLSKKFLKQDTYDV 617
Db 608 NPQPHIITILDPDLPNTSPFTAELTHGASVNWNTIEYNDAQESLILQPRKOLEIGEYKI 667
QY 618 HLSLDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGF-----ILPVLGAVLALLF 668
Db 668 HLKLDNQNDQVTTLDVHVCDCEGTNNC---MKAGIVAAGLQVPAIILGILGILLALI 724
QY 669 LLLVLLLVKRRKIKPELILPEDDTRDNVYVYEGEGGEEDQDYITOLHRLGLEARPEV 728
Db 725 LILLLLFLLRRRTVVEKPLPPDDTDRDNVYVYDEEGGEEDQDFLSQLHRLGLDARPEV 784
QY 729 VLRNDVAPTIPTPMYRPRPANPEIGNFTIENLKAANTDPTAPPYDTLLVDFYEGSGSD 788
```

Db 785 T-RNDVAPTLMSVQYRPRANPDEIGNFIDENLKADSDPTAPPYDLSLVDFYEGSGSE 843
Qy 789 AASLSSTSSASDQDQDYDLNENWGSFCKLADMYGGEDD 829
Db 844 AASLSLSNSESQDQDYDLNENWGNRFCKLADMYGGEDD 884

RESULT 7
US-08-237-919-2
; Sequence 2, Application US/08237919
; Patent No. 5610281
; GENERAL INFORMATION:
; APPLICANT: Brenner, Michael B
; APPLICANT: Cepek, Karyn L
; TITLE OF INVENTION: Methods and Compositions for
; Modulating Heterotypic E-cadherin Interactions with T lymphocytes
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,919
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 878 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-237-919-2

Query Match 53.8%; Score 2348.5; DB 1; Length 878;
Best Local Similarity 54.0%; Pred. No. 8.4e-210;
Matches 472; Conservative 115; Mismatches 230; Indels 57; Gaps 12;

Qy 8 LASLILLQVCMQLQAASPCRAVFEAEVITLEAGAEQEPQALGKV-FMCGPQGE--- 62
Db 8 LSGLLLLRSLPGSQERSPPCLRELHVGAPAPPEKRP--RLGRVNFEDCTGRQRTAI 65
Qy 63 -----PALFSTNDNDDFTVRNGETVQERSLK-----ERNP--- 92
Db 66 FLTPPIKVGTDGVTIVTKRPLRPHNPTDPFLGLRWDSTYRKFTKVTLLNTVGHHRPPPHQ 125
Qy 93 -----LKIIFP-SKEILRRHKEDWVAPISVPENCKGPPQRLNOLKSKDRTKIF 142
Db 126 ASVSGIQELITFPNNSPGLRRQRKDWVIPPISCENENKGPFPKNLVQIRKNKQKEGVF 185
Qy 143 YSITPGADSPPEGVAFVEKETGMLLNKPLDRREIAKYELFGHAVSNGASVEDPMNIS 202
Db 186 YSITQGGADTPVGVFIETERTGWLKVTEPLDRERIATYTLFSAVSSNGNAVEDPMEL 245
Qy 203 IIVTDQNDHKPKFTQDTRFGSVLEGLVPGTSVMQVATDEDAIYTYNGVVAISHQEP 262
Db 246 ITVTQDNDNKPEFTQEVFKGSMGALPGTSVMEVATDADDVNTYNAAIATYILSQDP 305

Qy 263 KDPHDLMTFTHRTGTSTISVSGLDREKVPYETLTQATMDGSGSTTTTAVAVVEILDAN 322
Db 306 ELDPDKMFINRNTGIVSVTTTGLDRESPTTLLVQAADLQEGSLSTTATVITVTDN 365
Qy 323 DNAPMTDPQKYEAHVPENAVGVQRLTVTDLDPNSPAWRATYILNGGGDGHFTITTH 382
Db 366 DNPPIFNPTTYKQVPEANEANVVITLKVTDADAPNTPAWEAVYTL-NDGGQGVVVTN 424
Qy 383 PESNOGILTRKGLDFEAKNQHTLYXEVTNEAPFVLKLPSTSTATIVVHVVDVNEAPVFP 442
Db 425 PVNNDGILKTAKGLDFEAKQYILHVAIVNVVPEVSLTSTATVTVDLVDVNEGPVFP 484
Qy 443 PSKVEVQEGIPTGEPVCVYTAEDPPK-ENQKISYRILRDPAGWLAMPDPSQGVAVGTL 501
Db 485 PEKRVESDEDFGVGQEITSYTAQEPDTFMEQKITRYRWRDTRNWLINPDTGAISRAEL 544
Qy 502 DRDEQFVRNIIYEVWVLMNDGSPPTGTGTLILLILIDVNDHGPVPEPQITICNQSPV 561
Db 545 DREDFEHVKNSTYALIIATDNGSPVATGTGTLILLILSDVNDNAPIPEPRTIFFCERNPK 604
Qy 562 RVLNITDKDLSPTSPFOAQLTDDSDIYVTAENE-EGDTVVLKFLKQDTYDVHLS 620
Db 605 PQVINIHADLPPTSPPTAEIETHGRVPNWTIQVNDPTQESIILKPKMALEVGDYKINLK 664
Qy 621 LSDHGNKEQLTVIRATVCDCHGHVETC--PGPMKGGF---ILPVLGAVLALLFLLLVLL 674
Db 665 LMDNQNKDQVTTLEVSVCDCGAAGVCRKAQPVVAGIQAIPAILGILGILALLLILL 724
Qy 675 LLVRKKRKIKEPILLBEDDTRDNVYFYGEGGEGEDDYDITQLHRLGLEARPEVVLNDV 734
Db 725 LFLRRRAVVKPELLPPEDDTRDNVYFYGEGGEGEDDYDITQLHRLGLEARPEV-NDV 783
Qy 735 APTIIPTPMYRPRANPDEIGNFIENLKAANTDPTAPPYDTLVDFYEGSGSAAASLS 794
Db 784 APTLMSVRYLPRANPDEIGNFIDENLKADTPTAPPYDLSLVDFYEGSGSAAASLS 843
Qy 795 LTSASDQDQDYDLNENWGSFCKLADMYGGED 828
Db 844 LNSSESQDQDYDLNENWGNRFCKLADMYGGED 877

RESULT 8
US-08-732-429-2
; Sequence 2, Application US/08732429
; Patent No. 6300080
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Methods and Compositions for Modulating Heterotypic
; E-cadherin Interactions with T Lymphocytes
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/732,429
; FILING DATE: herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/237,919
; FILING DATE: 3 May 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7023
; TELECOMMUNICATION INFORMATION:


```
Db 605 PQVINIHADLPNTSPFTAEATHGRVPMNTIQNDPTQBSIILKPKMALEVGDKINLK 664
Qy 621 LSDHGKEQLTVIRATVCDCHGHVETC--PGPWKGGF-----ILPVLGAVLALLFLIALL 674
Db 665 LMDNQNDQVTTLEVSVCDCGAGVCRKAQPVAGLQIPAILGILGAILLALLIALL 724
Qy 675 LLVRKKEIKPELLPEDDTRDNVYFYEGEGGEDQDYDITQLHRLGLEARPEVVLNDV 734
Db 725 LFLRRRAVKEPLLPEDDTRDNVYFYEGEGGEDQDFLSQLHRLGLDARPEVT-RNDV 783
Qy 735 APTIIPMYRPRANPDEIGNFIENLKAANTDPTAPPYDITLLVDFYEGSGDAASLSS 794
Db 784 APTLMSVPRYLPRANPDEIGNFIENLKAADTPTAPPYDLSLVDFYEGSGEASLSS 843
Qy 795 LTSASDQDDYDYLNEGWSRPFKLADMYGGGD 828
Db 844 LNSSESDQDDYDYLNEGWSRPFKLADMYGGGD 877

RESULT 10
US-09-798-267-3
; Sequence 3, Application US/09798267
; Patent No. 6406870
; GENERAL INFORMATION:
; APPLICANT: Brenner, Michael
; APPLICANT: Cepek, Karyn
; TITLE OF INVENTION: Methods and Compositions for Modulating Heterotypic E-cadherin
; FILE OF INVENTION: Interactions with T Lymphocytes
; FILE REFERENCE: L0560/7008ERP
; CURRENT APPLICATION NUMBER: US/09/798,267
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 08/237,919
; PRIOR FILING DATE: 1994-05-03
; PRIOR APPLICATION NUMBER: PCT/US 95/05518
; PRIOR FILING DATE: 1995-05-03
; PRIOR APPLICATION NUMBER: US 08/732,429
; PRIOR FILING DATE: 1996-11-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 878
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-798-267-3

Query Match 53.8%; Score 2348.5; DB 2; Length 878;
Best Local Similarity 54.0%; Pred. No. 8.4e-210;
Matches 472; Conservative 115; Mismatches 230; Indels 57; Gaps 12;

Qy 8 LASLILLOVCHLQCAASEPCRAVREAEVTLLEAGAGQEPQALGV-FWGCPCQE----- 62
Db 8 LSGLLLLRSPLGQERSPPCLTRHLVHGAPPEKRP--RLGRVNFEDCTGRQRTAI 65
Qy 63 -----PALFSTDNDDFTVRNGETVQERSLSK-----ERNP--- 92
Db 66 FLTPIPKVGTGVTIVKRLPHNPTDFLGLRWDSIYKFKSTVLTNVTGHHRRPPPHQ 125
Qy 93 -----LKIFP-SKRILRRHRKDWVAPISVPENKGPFPQRLNQLKSNKDRDRTKIF 142
Db 126 ASVSGIAELLTTFNNSPGLRQRKRDWVIPSICPENEGKGFPPKLVQIKSNKDKGKVF 185
Qy 143 YSITGPGADSPPGFAVEKETGMLLNKPLDRREIAKYELFHAVSENGASVEDPMNIS 202
Db 186 YSITGQGGADTFPVGVFIETRTGMLKVTEPLDRERIATYTLFHAVSSNGNAVEDPMEL 245
Qy 203 IIVTDQNDHKPKFQDTPFRGSLVLEGLVPCTSVMQVATDEDAIYTVNGVAVYIHSQEP 262
Db 246 ITVTDQNDKPEFTQEVFKGSMGALFCTSVMEVATDADDVNTYNAIATYILSQDP 305
Qy 263 KPDHLMFTIHRSTGTSIVISGLDREKVPYTTIQTATMDMDGSGSTTAVAVVEILDAN 322
Db 306 ELPDKNMTINRNTGVISVWTTGLDRSFPTTYTLVWQAADLQGBGLSTTATAVITVDTN 365
```

```
Qy 323 DNAPMFDPQYEAHVPENAVGHEVQRLVTDLDPNPSPAWRATYILIMGGDGDHFTITTH 382
Db 366 DNPPIFNPTTYKQVENEANVITLTKVTDADAPNTPAWEAVYTIL-NDGGQGVVTTN 424
Qy 383 PESNOGILTRKGLDPEAKNQHTLYVEVNEAFVLKLPSTSTATIVVHVEDVNEAPVFP 442
Db 425 PVNNDGLTKAGLDPEAKQOYLHVAVTNVVPEVSLTSTATVTVLDVNEGPVFP 484
Qy 443 PSKVEVQEGIPGEPVCVVYTAEDPK-ENOKLSYRILRDPAGWLAMDPSGOVAVGTL 501
Db 485 PEKREVESEDFGVQBSITSYAQEPDTFMEQKITIRIWRDRNWLEINPTGALSTRAEL 544
Qy 502 DRDEQFVRNIVYEVWVLAMDNGSPPTGTGTLTLLTLDVNDHGPVPEPRQITICNOSP 561
Db 545 DREDFEHVKNSTYALIATDNGSPVATGTGTLTLLSVDNDNAPIPEPTIIPFCERNPK 604
Qy 562 RHLVNTDMDLSPTSPFQAQLTDDSDIYVTAEVNE-EGTVTVLSLKKFKQDTYDVHLS 620
Db 605 PQVINIHADLPNTSPFTAEATHGRVPMNTIQNDPTQBSIILKPKMALEVGDKINLK 664
Qy 621 LSDHGKEQLTVIRATVCDCHGHVETC--PGPWKGGF-----ILPVLGAVLALLFLIALL 674
Db 665 LMDNQNDQVTTLEVSVCDCGAGVCRKAQPVAGLQIPAILGILGAILLALLIALL 724
Qy 675 LLVRKKEIKPELLPEDDTRDNVYFYEGEGGEDQDYDITQLHRLGLEARPEVVLNDV 734
Db 725 LFLRRRAVKEPLLPEDDTRDNVYFYEGEGGEDQDFLSQLHRLGLDARPEVT-RNDV 783
Qy 735 APTIIPMYRPRANPDEIGNFIENLKAANTDPTAPPYDITLLVDFYEGSGDAASLSS 794
Db 784 APTLMSVPRYLPRANPDEIGNFIENLKAADTPTAPPYDLSLVDFYEGSGEASLSS 843
Qy 795 LTSASDQDDYDYLNEGWSRPFKLADMYGGGD 828
Db 844 LNSSESDQDDYDYLNEGWSRPFKLADMYGGGD 877

RESULT 11
PCT-US95-05518-2
; Sequence 2, Application PC/TUS9505518
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Methods and Compositions for Modulating Heterotypic
; TITLE OF INVENTION: E-cadherin Interactions with T Lymphocytes
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05518
; FILING DATE: herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/237,919
; FILING DATE: 3 May 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
```



```

Db      364 NTATAVITVTDVNDNPPEFTAMTFYGEVPEPNRVDIIIVANLTVTDKQOHPTPAWNNAVYRIS 423
Qy      370 GGDGDHFTITTHPESNOGLITRKGLDPEAKQOHTLYVEVTNEAPFVLXL---PTSTAT 426
Db      424 GGDPTGRFAJQTDNSNDGLVTVVKPIDFETNRMFVLTVAAENQVPLAKGIOHPPQSTAT 483
Qy      427 IVHVEDVNEAPFVPPKVVVEQGIPTGEPCVCTAEDPK-ENQKISYRIILRDPAGW 485
Db      484 VSVTVIDVNEPYFAPNPKIIRQEEGLHAGTMLTFTTAQDPDRYMQQNIYTKLSDPANW 543
Qy      486 LAMPDQSGQVAVGTLREDEQFVRNNIYVMVLAMNGSPPTGTGTLTLLTLIDVNDHG 545
Db      544 LKIDPVNGQITTIIVLDRESN-VKNNIYNATFLASDNGIPPMSCGTGLQIYLLIDINDNA 602
Qy      546 PVPEPRQITICNOSPVRHVNIT--DKDLSPHTSFQOQLTDDSDIY---WTAEVNEEGD 600
Db      603 FQVLPQEAETC-ETPDNSINITALDYDIDFNAGFFAFDLPLSPVTIKRNWTI--TRLNGD 660
Qy      601 TVVLSLK-KFLKQDQTYVHLSLSDHGN--KEQLTVIRATVCDCHG-----VETCPGPMK 652
Db      661 FAQLNLKIKLEAGIYEVPIIITDSGNPPKSNISILAVKYQCQDSNGDCTDVRIVGAGL 720
Qy      653 GGFILPVLGAVLALLFLLVLLLV-----RKRRKIKEPLLLPEDDTRDNVYYG 702
Db      721 G-----TGATIAILLGIIILLILVLMFVVMKRRDKERQAKQLLIDPEDDVRDNILKYD 774
Qy      703 BEGGGEEDQDYDITQLHGLEARPEV-----VLRNDVAPTIIPTPMYRPRPA--NPDEIG 755
Db      775 BEGGGEEDQDYDLSLOQOPDTVEPDAIKPVGIRMDERP-IHAEPQYFVRSAAHPGDI 833
Qy      756 NFIIENLKAANTDPTAPPYDTLVFDYEGSGSDAASLSLTSSASDQDQDYDLNWSGR 815
Db      834 DFINEGLKAANDPTAPPYDLSLVFDYEGSGSTAGSLSSLNSSSGGEQDYDYLDNDGPR 893
Qy      816 FKLADMYGGGED 828
Db      894 FKLADMYGGGDD 906

```

Search completed: December 12, 2005, 09:17:05
Job time : 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2005, 09:11:23 ; Search time 11 Seconds
(without alignments)
420.863 Million cell updates/sec

Title: US-10-696-639-39
Perfect score: 4369
Sequence: 1 MGLPRGPLASLLLLQVCMLO.....NEWGRFKLADMYGGGEDD 829

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 32527 seqs, 5584426 residues

Total number of hits satisfying chosen parameters: 32527

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA New.*
- 1: /cgn2_6/prodata1/pubpa/US09_NEW_PUB.pap.*
 - 2: /cgn2_6/prodata1/pubpa/US06_NEW_PUB.pap.*
 - 3: /cgn2_6/prodata1/pubpa/US07_NEW_PUB.pap.*
 - 4: /cgn2_6/prodata1/pubpa/US08_NEW_PUB.pap.*
 - 5: /cgn2_6/prodata1/pubpa/PCT_NEW_PUB.pap.*
 - 6: /cgn2_6/prodata1/pubpa/US10_NEW_PUB.pap.*
 - 7: /cgn2_6/prodata1/pubpa/US11_NEW_PUB.pap.*
 - 8: /cgn2_6/prodata1/pubpa/US60_NEW_PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4361	99.8	829	7	US-11-090-739-122
2	4361	99.8	829	7	US-11-186-284-18
3	982	22.5	790	7	US-11-080-991-82
4	957	21.9	896	7	US-11-080-991-98
5	919	21.0	801	7	US-11-174-150-29
6	796	18.2	824	6	US-10-821-234-1008
7	586	13.5	589	7	US-11-174-150-28
8	526.5	12.1	832	7	US-11-108-172-1081
9	526.5	12.1	958	7	US-11-108-172-1087
10	408.5	9.3	1060	7	US-11-090-739-120
11	381.5	8.7	807	7	US-11-102-240-98
12	333.5	7.6	1184	6	US-10-131-826A-394
13	280	6.4	530	6	US-10-131-826A-130
14	240.5	5.5	1730	7	US-11-192-967-4
15	240.5	5.5	1730	7	US-11-193-715-4
16	226	5.2	72	6	US-10-509-170-19
17	214	4.9	1734	7	US-11-192-967-6
18	214	4.9	1734	7	US-11-193-715-6
19	207	4.7	1717	7	US-11-192-967-2
20	207	4.7	1717	7	US-11-193-715-2
21	147.5	3.4	2384	6	US-10-821-234-1545
22	126.5	2.9	1346	7	US-11-060-005-2
23	125.5	2.9	2647	6	US-10-821-234-1303
24	124.5	2.8	895	7	US-11-150-406-2
25	124.5	2.8	920	6	US-10-821-234-1129

26	123.5	2.8	826	6	US-10-793-626-1066	Sequence 1066, Ap
27	121	2.8	428	6	US-10-689-742-50	Sequence 50, Appl
28	118.5	2.7	559	6	US-10-793-626-1376	Sequence 1376, Ap
29	118	2.7	1596	7	US-11-060-005-4	Sequence 4, Appli
30	107.5	2.5	616	7	US-11-090-878-20	Sequence 20, Appl
31	107.5	2.5	624	7	US-11-090-878-2	Sequence 2, Appli
32	107.5	2.5	637	7	US-11-090-878-24	Sequence 24, Appl
33	106.5	2.4	596	7	US-11-090-878-22	Sequence 22, Appl
34	106.5	2.4	617	7	US-11-090-878-18	Sequence 18, Appl
35	106.5	2.4	3063	7	US-11-186-284-26	Sequence 26, Appl
36	105.5	2.4	5024	6	US-10-793-626-2964	Sequence 2964, Ap
37	105	2.4	720	6	US-10-793-626-2058	Sequence 2058, Ap
38	105	2.4	1897	6	US-10-821-234-1635	Sequence 1635, Ap
39	105	2.4	1907	7	US-11-000-463-250	Sequence 250, App
40	104	2.4	7968	7	US-11-186-731-5	Sequence 5, Appli
41	103.5	2.4	598	7	US-11-090-878-8	Sequence 8, Appli
42	103.5	2.4	910	6	US-10-131-826A-112	Sequence 112, App
43	103.5	2.4	1141	6	US-10-601-368-24	Sequence 24, Appl
44	103.5	2.4	1166	6	US-10-601-368-22	Sequence 22, Appl
45	103.5	2.4	1188	6	US-10-601-368-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-11-090-739-122
; Sequence 122, Application US/11090739
; Publication No. US20050260639A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Yusuke
; APPLICANT: KATAGIRI, Toyomasa
; APPLICANT: NAKAGAWA, Hirowaki
; TITLE OF INVENTION: METHOD FOR DIAGNOSING PANCREATIC CANCER
; FILE REFERENCE: 082368-003600US
; CURRENT APPLICATION NUMBER: US/11/090,739
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: PCT/JP2003/011817
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: US 60/555,809
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: US 60/450,889
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/414,872
; PRIOR FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 122
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-739-122

Query Match 99.8%; Score 4361; DB 7; Length 829;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 828; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MGLPRGPLASLLLLQVCMLOCAASPCRAVFREAEVTL	EAGAEQEPGQALGVFMGCPG	60
Db	1	MGLPRGPLASLLLLQVCMLOCAASPCRAVFREAEVTL	EAGAEQEPGQALGVFMGCPG	60
Qy	61	QSPALFSTDDNDFTVRNGETVQERRSLKERNPLKIPPSKRI	LRHRKRDVWVAPISVPENG	120
Db	61	QSPALFSTDDNDFTVRNGETVQERRSLKERNPLKIPPSKRI	LRHRKRDVWVAPISVPENG	120
Qy	121	KGPPFQRLNQLKSNKDRDTKIFYSITGCGADSPPGVFAVEK	ETGWLILNKPLDREIEIAK	180
Db	121	KGPPFQRLNQLKSNKDRDTKIFYSITGCGADSPPGVFAVEK	ETGWLILNKPLDREIEIAK	180
Qy	181	YELFGHVSNGASVEDPMNISIIITDQNDHHPKPTQDTFRGS	VLGVLPGTSVMQVTAT	240
Db	181	YELFGHVSNGASVEDPMNISIIITDQNDHHPKPTQDTFRGS	VLGVLPGTSVMQVTAT	240

```
Qy 241 DEDDAIYTYNGVAVSIHQSPKDPHLMFTIHRSTGTISVSSGLDREKVPYTLTQA 300
Db 241 DEDDAIYTYNGVAVSIHQSPKDPHLMFTIHRSTGTISVSSGLDREKVPYTLTQA 300
Qy 301 TMDGSGSTTAVAVBEILDANDNAPMDFPKQYEAHVPENAVGHVQRLTVTDLADPNSP 360
Db 301 TMDGSGSTTAVAVBEILDANDNAPMDFPKQYEAHVPENAVGHVQRLTVTDLADPNSP 360
Qy 361 AWRATYLMGDDGDHFTIITHPSNQGLTTRKGLDFEAKNQHTLVVEVNEAPFVLKL 420
Db 361 AWRATYLMGDDGDHFTIITHPSNQGLTTRKGLDFEAKNQHTLVVEVNEAPFVLKL 420
Qy 421 PTSTATIIVHVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDDPKENQKISYRILR 480
Db 421 PTSTATIIVHVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDDPKENQKISYRILR 480
Qy 481 DPAGWLAMPDSGQVTAAGTLTDRDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLLID 540
Db 481 DPAGWLAMPDSGQVTAAGTLTDRDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLLID 540
Qy 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSLKKFLKQDYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILPVL 660
Db 601 TVVLSLKKFLKQDYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILPVL 660
Qy 661 GAVLALLFLLVLLVLRKKRKIKEPILLPDDTRDNVFFYGGEGGEEDQDYDITQLHR 720
Db 661 GAVLALLFLLVLLVLRKKRKIKEPILLPDDTRDNVFFYGGEGGEEDQDYDITQLHR 720
Qy 721 GLEARPEVLRNDVAPTIIPPMYRPRPANDPDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVLRNDVAPTIIPPMYRPRPANDPDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
Qy 781 DYEAGSDAASLSLTSASDQDQDYDLNENWGSFRFKKLADMYGGGDD 829
Db 781 DYEAGSDAASLSLTSASDQDQDYDLNENWGSFRFKKLADMYGGGDD 829
```

RESULT 2

```
US-11-186-284-18
; Sequence 18, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kanatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MP01-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; PRIOR FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 829
; TYPE: PRT
```

```
; ORGANISM: Homo Sapiens
US-11-186-284-18
Query Match 99.8%; Score 4361; DB 7; Length 829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 828; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MGLPRGPLASLLQLQWLCQAASEPCRAVFRAEVTLEAGAEQEPGQALGVFMGCPG 60
Db 1 MGLPRGPLASLLQLQWLCQAASEPCRAVFRAEVTLEAGAEQEPGQALGVFMGCPG 60
Qy 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDWVAPISVPENG 120
Db 61 QEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIPPSKRILRRHKRDWVAPISVPENG 120
Qy 121 KGFPPORLNOLKSNKORDTKIFYSIITGPGADSPPEGVFAVEKETGMLLNKPLDREIAK 180
Db 121 KGFPPORLNOLKSNKORDTKIFYSIITGPGADSPPEGVFAVEKETGMLLNKPLDREIAK 180
Qy 181 YELFGHAVSENGASVEDPMNISIITDONDHKPKFTQDTPRGVLEGLPCTSVMOVTAT 240
Db 181 YELFGHAVSENGASVEDPMNISIITDONDHKPKFTQDTPRGVLEGLPCTSVMOVTAT 240
Qy 241 DEDDAIYTYNGVAVSIHQSPKDPHLMFTIHRSTGTISVSSGLDREKVPYTLTQA 300
Db 241 DEDDAIYTYNGVAVSIHQSPKDPHLMFTIHRSTGTISVSSGLDREKVPYTLTQA 300
Qy 301 TMDGSGSTTAVAVBEILDANDNAPMDFPKQYEAHVPENAVGHVQRLTVTDLADPNSP 360
Db 301 TMDGSGSTTAVAVBEILDANDNAPMDFPKQYEAHVPENAVGHVQRLTVTDLADPNSP 360
Qy 361 AWRATYLMGDDGDHFTIITHPSNQGLTTRKGLDFEAKNQHTLVVEVNEAPFVLKL 420
Db 361 AWRATYLMGDDGDHFTIITHPSNQGLTTRKGLDFEAKNQHTLVVEVNEAPFVLKL 420
Qy 421 PTSTATIIVHVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDDPKENQKISYRILR 480
Db 421 PTSTATIIVHVEDVNEAPFVPPSKVVEQGIPTGEPVCVYTAEDDPKENQKISYRILR 480
Qy 481 DPAGWLAMPDSGQVTAAGTLTDRDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLLID 540
Db 481 DPAGWLAMPDSGQVTAAGTLTDRDEQFVRNNIYEVNVLAMDNQSPPTTGTGTLTLLID 540
Qy 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Db 541 VNDHGPVPEPQITICNQSPVRHLNITDKDLSPTSPFOAQLTDDSDIYWTAEVNEEGD 600
Qy 601 TVVLSLKKFLKQDYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILPVL 660
Db 601 TVVLSLKKFLKQDYDVHLSLSDHGNKEQLTVIRATVCDCHGHVETCPGPKWGGFILPVL 660
Qy 661 GAVLALLFLLVLLVLRKKRKIKEPILLPDDTRDNVFFYGGEGGEEDQDYDITQLHR 720
Db 661 GAVLALLFLLVLLVLRKKRKIKEPILLPDDTRDNVFFYGGEGGEEDQDYDITQLHR 720
Qy 721 GLEARPEVLRNDVAPTIIPPMYRPRPANDPDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
Db 721 GLEARPEVLRNDVAPTIIPPMYRPRPANDPDEIGNFIIENLKAANTDPTAPPYDTLLVF 780
Qy 781 DYEAGSDAASLSLTSASDQDQDYDLNENWGSFRFKKLADMYGGGDD 829
Db 781 DYEAGSDAASLSLTSASDQDQDYDLNENWGSFRFKKLADMYGGGDD 829
```

RESULT 3

```
US-11-080-991-82
; Sequence 82, Application US/11080991
; Publication No. US20050266437A1
; GENERAL INFORMATION:
; APPLICANT: Velby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
```



```
Db 714 VFGATKGR-----FPEDIAQONLIISNTEAFGDDRVCSANGFMTQTNNSSQGFPGT 766
Qy 726 PEVLRNDVAPTIIPTMYRPRANPDEI---GNFIENLKAAN-----766
Db 767 MSGMKNGGQETI-----EMKGGNQTLSCRGAGHHHTLDCRGGHTEVD 812
Qy 767 -----TDP-----TAPPYDTLLVFDYEGSGDAASLSLTSAS 800
Db 813 NCRITYSEWSFTOPRLGEKLRHCNQNEDRMPSQDYVLTYNREGSGPAGSVG---CCSEK 870
Qy 801 DQDQDYDYLNEWGRFRKKLAD 821
Db 871 QEEDGLDFLNNLEPKFITLAE 891

RESULT 5
US-11-174-150-29
; Sequence 29, Application US/11174150
; Publication No. US20050260714A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall F.
; APPLICANT: Xiang, ZhaoYing
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50022
; CURRENT APPLICATION NUMBER: US/11/174,150
; CURRENT FILING DATE: 2005-07-01
; PRIOR APPLICATION NUMBER: US/10/257,174
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: PCT/US01/11797
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/196,603
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/199,417
; PRIOR FILING DATE: 2000-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-174-150-29

Query Match 21.0%; Score 919; DB 7; Length 801;
Best Local Similarity 33.5%; Pred. No. 3.8e-62;
Matches 255; Conservative 128; Mismatches 322; Indels 56; Gaps 24;

Qy 103 RHHRDWTVPISVPENKGPFPQRLNOLKSNKDR-DTKIFYSITGCGADSPPEGFAVE 161
Db 55 QTRKRWNNQFFVLEEYTGTDPLVYGLKHSMDRGDSIKYILSGEGAGI-----VFTID 110
Qy 162 KETGWLINLKPDLREEIAKELFGHAV--ENGASVEDPMNISIIIVTDQNDHKPKFTQDTF 220
Db 111 DTTGDIHAIQLDREERAQYTLRAQALDRRTGRPEPESEFIKIQTIDNNEPKFLDGPY 170
Qy 221 RGSVLEGLVPGTSVMQVATDEDAIYYTNGVWVAYS IHSQBPKPDPHLMFTIHRSTGTIS 280
Db 171 VATVPMSFVGTSVIQVATDADDPYTGNSARVVSILQGP-----YFSVDSKGTGVR 224
Qy 281 VISSGLDREKPEYTLTIQATMDGD--GSTTTAVAVVEIILDANDNAMPDPQKYEAVHP 338
Db 225 TALMMMDREAKYEYEVIIQAKDMGQLGLAGTTVNTILSDVNNDPPRPFQKHQMSVL 284
Qy 339 ENA-VGHEVQRILTVDLAPNSPAWRATYILIMGDGDGDHFTITTHPESNQILTRKGLD 397
Db 285 ESAPISSTVGRVFAKDLDEGINAEK--YTIVDGADAFDIDSTDPNFQVGIITVKKFLS 342
Qy 398 PEAKNQHTLYVEVTN---EAPFVLKLP--TSTATIVVHVVDNEAPVFPVPSKVEVQEGI 453
Db 343 FESKKSYYLTKVEGANPHLEMRFLNLGFPQDTTTHVISVEDVDPEVPFPGFYFVEVPEDV 402
```

```
Qy 454 PTGEPVCVYTAEDPDKENOKISYRIIL--DPAGLWAMPDSDGQVAVGTLDREDEQFVRN 511
Db 403 AICTTIIISAKDPDVTNNNSIRYSIDRSSDPGRFFVYVDITTGALMTARPLDREFSW--- 459
Qy 512 NIYEVWVLAMDNQSPPTTGTGTLTLLIIVNDHGP-VPEPRQITICNQSPVRHVL---NI 567
Db 460 --HNITVTLAMENMNPQVGSVPVTIKVLDVNDNAPEFPFRFEYAFVCENAKAGOLIQTVSA 517
Qy 568 TKDLSPHSTPFOAQLTDD--SDIYTAEVNEBGDTVVLSSKKFLKQDQTDV--HLS--LS 622
Db 518 VDODDPRNGOHFYSLAPEAANNPNFTIRDQNTARILTRRSGRFOEQSVFHPILIIA 577
Qy 623 DHG-----NKEQLTVIRATVCDCHGVETCTCPGWKGFILPV-----LGAVLALLFLLV 672
Db 578 DSGQPVLSSTGTLTIQVCSDDDDGHVMSCS--PEAYMLPVLSLRSRGALIIAILACIFVLV 634
Qy 673 LLLLV-RKKRKIKEPULL-PEDTRONVIFYEGEGEEDQD-YDITQLHRGLEAREPV 729
Db 635 LVLLILSMRRHRKQPYIIDDEENIHENIVRYDDGEGEEDTEAFDIAAMWNPREAQAGAA 694
Qy 730 --LRNDVAPTIIPTMYRPR-ANPDEIGNFIENLKAANTDPTAPPYDTLLVFDYEGSG 786
Db 695 PKTRQDMLPEIESLSRYVPOTCAVNSTVHSYVLAKLYEADMDLWAPFDSLOTVMFGDG 754
Qy 787 SDAASISLTSASDQDDQDYDYLNEWGRFRKKLADMYGGE 827
Db 755 SVAGSLSSLOSATSSEQSFDLTDMGPRPRKLAELYGASE 795

RESULT 6
US-10-821-234-1008
; Sequence 1008, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeciampaia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: Pct_SEQ_genes Version 1.0
; SEQ ID NO 1008
; LENGTH: 824
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1008

Query Match 18.2%; Score 796; DB 6; Length 824;
Best Local Similarity 32.1%; Pred. No. 8.2e-53;
Matches 251; Conservative 124; Mismatches 323; Indels 84; Gaps 33;

Qy 89 ERNPLKIFPSKRILRRHKRDWVAPISVPENKGPFPQRLNOLKSNKDRDKIFVSTGP 148
Db 72 QRDTHSLPLPH---RRORRWIWNQMIHDEEKNTSLPHHVGVKIKSSVSR-KNAYLLKGE 127
Qy 149 GADSPPEGVFAVEKETGWLINLKPDLREEIAKELFGHAV--SENGASVEDPMNISIIIVTD 207
Db 128 YVGK---VFRVDAETGDFAIERLDRDRENISEYHLTAVIVDKOTGENLETSPSTIKVHD 183
Qy 208 QNDHKPKFTQDTFRGSVLEGLVPGTSVMQVATDEDAIYYTNGVWVAYS IHSQBPKPDPH 267
Db 184 VNDNWPVFTHRLENASVPSSAVGTSVISTAVDADDPYTGDSHASVMYQILKGE----- 238
Qy 268 LMFTIHSTGTISVISSGLDREKPEYTLTIQATMDG-DGSTTTAVAVVEIILDANDNAP 326
Db 239 -YFALDINS-GRITITISLDRKQARVEIIVVEARDAQGLRDSGTATVLTQLQINDNFP 296
Qy 327 MFDQPKYEAHPENA-VGHEVQRILTVDLAPNSPAWRAT-YLIMGDGDGDHFTITTHPE 384
```

Db 297 FFTQTKYTFVVPEDTRVGTSGVSLFVED---PDEPQNRMTKYSLRGDYQDAFTIETNFA 353
Qy 385 SNOGILTRKGLDFEAKNQHTLVVEVTN---EAPFVLKLPSTSTATVWHVEDVNEAPVFW 441
Db 354 HNEGILKPKPLDYEQYQSFIVEATDPTIDIRYMSPPAGNRAQVLIINITDVEDEPIQ 413
Qy 442 PPSKWEVQSGITGPEP-VCVYTAEDPKENQKISYRILRDPAGWLAMPDSCQ---VTA 497
Db 414 QPFYHQLKEN--QKKPLIGTVLAMPDPAARHSIGYSIRR-----TSDKGQFFRVTK 463
Qy 498 VGTLDREDEQFVRNNIYEVVWVLANMGSPPTTGTGT-----LLLTLDVNDHGP-VPE 549
Db 464 KGDYINEKE--LDREVYPMYNLTVAEKELDSTGTPTGKESIVQVHIEVLDENDNAPEFAK 521
Qy 550 PROITTC-NQSPRVHVLNIT--DKDLSPTSPOAOLTDSDIYMTAEVNEEGDTVVLSL 606
Db 522 PYQPKVCENAVHQLVLIQSAIDKDIYPRNVKFKLTNTENFTLT--DNHNTANITV 578
Qy 607 K--KFLKQDITYDVH---LSLSDHG--NKEQLTVIRATVCDCH--GHVETCP--GPWKGGF 655
Db 579 KYGFQDREHT-KVHFPLVVISDNGMPSRTGTSTLTVAVCKNCQGEFTFCEDMAAQGVVS 637
Qy 656 ILPVLGAVLALLFLLVLLVLRKKIKIKEPLI-----LPEDDTRDNVYFYGEGGGEED- 710
Db 638 IQAWAAILLCILITITVITLLIFLRRRLRKOARHGKSVPE--THEQLVTVYDEGGGEMDT 695
Qy 711 QDYDITOLH--RGLEARPEVLRNDVAPTIPTPMYRPPA-----NPDEIGNFIENL 762
Db 696 TSDVSVLNVRRGGAKPPRAL--DARPSLYAQVQKPPRHAGCAGGCGEAMAEVKK 753
Qy 763 KAANTDPTAPPYDTLAVFDYEGSGSAASSLTSSASDODDYDVLNENWGRFKKLADM 822
Db 754 DEADHDGDPYDTHLYHYEGSGESAESLSLGTSDSDSDVDYDFLNDNGPRFKVLAEL 813
Qy 823 YG 824
Db 814 YG 815

RESULT 7

US-11-174-150-28
; Sequence 28, Application US/11174150
; Publication No. US20050260714A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall P.
; APPLICANT: Xiang, Zhaoqing
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50022
; CURRENT APPLICATION NUMBER: US/11/174,150
; CURRENT FILING DATE: 2005-07-01
; PRIOR APPLICATION NUMBER: US/10/257,174
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: PCT/US01/11797
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/196,603
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/199,417
; PRIOR FILING DATE: 2000-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-174-150-28

Query Match 13.5%; Score 588; DB 7; Length 589;
Best Local Similarity 34.1%; Pred. No. 2.9e-37;
Matches 165; Conservative 72; Mismatches 215; Indels 32; Gaps 13;

Qy 103 RRHKRDWVAPISVPENGKGPFPQRLNQLKSNKDR-DTKIFYSITGPGADSPPEGVAVE 161
Db 55 QRTKRSWVWVQFVLEEYTGTDPLVYGLHSDMDRGDSIKYILSGEGAGI-----VFTID 110
Qy 162 KETGWLILNKPDLREIEAKVELFGHAVS-ENGASVEDPMNISIIVTDQNDHKPKFTQDTF 220
Db 111 DTTGDIHAIQRLDREERAQYTLRAQALDRRTGRPMPESEFEIKIQDINDNEPKFLDGPY 170
Qy 221 RGSVLEGLVPGTSVMQVATDDEDAIYTYGVVVAYSIHSQEPKDPHDLMTFTHRSCTGIS 280
Db 171 VATVEMSPVGTSVIQVTATDADDDPTYGNSARVYVSILOQP-----YFVSDSKTGVIR 224
Qy 281 VISSGLDREKVEPYTLTIQATDMGD--GSTTAVAVVEILDANDNAPMFPDPOKYEAHVP 338
Db 225 TALMNDREAKEVEYEVIIQAKDMGGQLGLAGTTTNNITLSDVNDNPPRFPQKHQMSVL 284
Qy 339 ENA-VGHEVQRLTVDLDAPNSPAWRATYLIIMGDDGDHFTTTHPESNQILTTKGLD 397
Db 285 ESAPISSTVGRVFAKDLDEGINAEMK--YTIVDGADAFDIDSTDFNFQVGIITVKKPLS 342
Qy 398 FEAKNQHTLYVEVTN---EAPFVLKLP-TSTATIVVHVEDVNEAPVFPVPPSKVVEQEGI 453
Db 343 FESKSYTLKVGANPHLEWRFLNLGFPQDPTTVHLSVEDVDVDEPPVFEFGFYFEVEPDEV 402
Qy 454 PTGEPYCVYTAEDPPDKENQKISYRILR--DPAGWLAMPDPSGQVTVAGTLDREDEQFVRN 511
Db 403 AIGTTTIIISAKDPDVTNNISYSDRSSDPGRFFVYDITTGALMTARPLDREEPSW--- 459
Qy 512 NIYEVWVLANMGSPPTTGTGTLTLLTLDVNDHGP-VPEPRQITICNQSPVRHL---NI 567
Db 460 --HNITVLAMENMNPVSGSVFVTIKVLDVNDNAPEFPRFYFAVCENAKAGOLIQTVA 517
Qy 568 TDKD 571
Db 518 VDQD 521

RESULT 8

US-11-108-172-1081
; Sequence 1081, Application US/111108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11/108,172
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 10/025,380
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 09/922,217
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/833,263
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 09/649,811
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 09/609,448

```

; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/575,251
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/519,444
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/504,629
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: US 09/480,321
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: US 09/476,296
; PRIOR FILING DATE: 1999-12-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1081
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-172-1081

Query Match      12.1%; Score 526.5; DB 7; Length 832;
Best Local Similarity 26.7%; Pred. No. 2.2e-32;
Matches 190; Conservative 109; Mismatches 308; Indels 105; Gaps 27;

Qy      59  PGOEPALF--STDNDDFTRNGETVQE-----RRSLKER 90
Db      141  PG-KPFLYVNATDLDPATNGQLYYQIVQLPMINNMVYFQINNKTGAISLTREGSOEL 199

Qy      91  NPLKIFPSKRLRHKR-----DWWVA-----PISVPENKGGPPQR 127
Db      200  NPAK-NPSYNLVISVKDMGGQSENFSFDTTSVDIIVTENIWKAPKPVEMVENSSTDHPHPIK 258

Qy      128  LNQLKSNKDRDTKIFYSTIGTGADSPPEGVFAVEKETGWLILNKPDLREELAKYELFGHA 187
Db      259  ITQVRWN---DPGAQYSLV--DKEKLPFRFPESIDQE--GDIYVTPQLDREEDKDAYVFAVA 312

Qy      188  VSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVATATDEDDAIY 247
Db      313  KDEVGKPLSYDLEIHVKVKDINDNPPTCPSPVTVFEQVENERLNSIGTLTAHDDREDE-N 371

Qy      248  TYNGVAYSIHQBPQPHDLMTFTHRSTGTIFSVISGLDREKPEYVTLTITQATMDQGDG 307
Db      372  TANGFLNRYIVEQTPKPLMDGLFLIQTVAGMLQAKQSLKKQDTPOYNLTITIEVSDKO--- 428

Qy      308  STTTAVAVVELDANDNAMPDPQKYE--AHVPENAVGHEVQRLTVTDLDPNAPSAPKAT 365
Db      429  FKTLCFVQINVIDNDQIPIFEKSDYGNLTLAEDTNIGSTILTITQATDADSPFTGSSKIL 488

Qy      366  YLIMGDDGDHFTTTHPESNOGILTRKGLDFAKNQHTLYVEVTNEAPPVLKL---PT 422
Db      489  YHIKGSSEGLGVDTDHTHTGVYIHKPLDFETAASVNIKFAENPEPLVFGVKYNAS 548

Qy      423  STATIVVHVEDVNEAPVFPSPKVVVEQEIPTGEPVCVYTAEDPDKENKISYRIILRDP 482
Db      549  SFAKFTLIVTDVNEAPQSQHVQAKVSEDVAIGTKVGNVTAKDP--EGLDISYSLRGDT 606

Qy      483  AGWLAMPDPSQVAVGTLDREDSQFVRNNIYEVWVLAMDNQSPPTGTGTLTLITLDVN 542
Db      607  RGWLKIDHTVGEIIFSVAFLDRE-----AGSPYRQVQVATEVGGSSLSVSEFHLILMDVN 661

Qy      543  DHGP--VPEPQITCN--QSPVRHLNITDKD---LSPH--TSPFOAQLTDDSDIYWT 592
Db      662  DNPRLAKDYTGLEFCHPLSPAGSLIFEAATDDQHLFRGPHFTFSLGSGSLQNDWEV--- 718

Qy      593  AEVNEEGTIVVLSLKKF--LKQDTYDVHLSLDHGNK--EQLTVIRATVCDCHGIVE--TCP 648
Db      719  SKIN--GTHARLSSTRHTDFEERAYVVLIRINDGGRPPLEGIVSLPVTFCSG---VEGSCF 773

Qy      649  GPWKGFTLPVLGAVLALLF--LLLVLLLLLVKKRKKIKEPELILLDEDDTRDNY 598
Db      774  RPAGHQTGPTVGMAVGILLTLLLVIGIILAVFIRIK-----KQKQKQNV 819
```

```

RESULT 9
US-11-108-172-1087
; Sequence 1087, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuguu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11/108,172
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 10/025,380
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 09/922,217
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/833,263
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 09/649,811
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 09/609,448
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/575,251
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/519,444
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/504,629
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: US 09/480,321
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: US 09/476,296
; PRIOR FILING DATE: 1999-12-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1087
; LENGTH: 958
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-172-1087

Query Match      12.1%; Score 526.5; DB 7; Length 958;
Best Local Similarity 26.7%; Pred. No. 2.7e-32;
Matches 190; Conservative 109; Mismatches 308; Indels 105; Gaps 27;

Qy      59  PGOEPALF--STDNDDFTRNGETVQE-----RRSLKER 90
Db      267  PG-KPFLYVNATDLDPATNGQLYYQIVQLPMINNMVYFQINNKTGAISLTREGSOEL 325

Qy      91  NPLKIFPSKRLRHKR-----DWWVA-----PISVPENKGGPPQR 127
Db      326  NPAK-NPSYNLVISVKDMGGQSENFSFDTTSVDIIVTENIWKAPKPVEMVENSSTDHPHPIK 384

Qy      128  LNQLKSNKDRDTKIFYSTIGTGADSPPEGVFAVEKETGWLILNKPDLREELAKYELFGHA 187
Db      385  ITQVRWN---DPGAQYSLV--DKEKLPFRFPESIDQE--GDIYVTPQLDREEDKDAYVFAVA 438

Qy      188  VSENGASVEDPMNISIIITDQNDHKPKFTQDTFRGSVLEGLVPGTSVMQVATATDEDDAIY 247
```


Sequence 98, Application US/11102240
Publication No. US20050260647A1
GENERAL INFORMATION:
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
TITLE OF INVENTION: ESOPHAGEAL TUMOR
FILE REFERENCE: P3230KIC106C
CURRENT APPLICATION NUMBER: US/11/102,240
CURRENT FILING DATE: 2005-04-08
PRIOR APPLICATION NUMBER: 10/063662
PRIOR FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 10/006867
PRIOR FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 199-12-09
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 98
LENGTH: 807
TYPE: PRT
ORGANISM: Homo Sapien
US-11-102-240-98

Query Match 8.7%; Score 381.5; DB 7; Length 807;
Best Local Similarity 20.9%; Pred. No. 2e-21;
Matches 170; Conservative 99; Mismatches 251; Indels 293; Gaps 25;

Qy 114 ISVPENKGPORLNOLKSKND-RDTKIFVSIITGPGADS--PPEGVFAVEKETGWLLEN 170
Db 27 VEVPENYGNPFLYTLKPLUPREGAQIVLUS-----GDSKGATEGPFAMDPDSGFLVLT 81

Qy 171 KPLDREBIAYELFCHAVSENGAGVEDPMNIIIVTDQNDHKPKFTQDTRFGSVLEGVLP 230
Db 82 RALDREAEQVQLQVLTLEMQDGHVLMGPQVLVHVHKENDQVPHFSQAIYRALSRTGP 141

Qy 231 GTSVMQVATDEDAIYTYNGVAYSISHPKDPHLMFTI----- 272
Db 142 GIPELFLEASDRDEP-GTANSDLRPHILSQAPAQSPDMFOLEPRLGALALSPKGSTSLD 200

Qy 273 -----HSTGTISV----- 281
Db 201 HALERTYQLLVQVKMDGQASGHQATATVEVSIITWVSLEPITHLAENLKVLYPHHMAQ 260

Qy 282 -----ISSGLDREKVPYETTLTIQATMDGDSGTTA 312
Db 261 VHWSGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPL 320

Qy 313 VAVVEILDANDNAPMFPQKYEAHYPE-NAVGHVEQRLTVTDLDPNPSPAWRATYLLIM-- 369
Db 321 ELHLVMDNDNVNPICPDRPTVSIPELSPGTEVTRLSAEDADAPGSPNSHVYQLLSP 380

Qy 370 -----CG----- 371
Db 381 EPEDVEGRAFOVDPTSGSVTLGVLPRAGQNILLVLAIDLAGEGFGSTCEVEAVT 440

Qy 372 DDGDH-----FT 378
Db 441 DINDHAPFITSQIGTISLPEDVBPGLVAMLTADADLEPAFLMDFAIERGDTGTFG 500

Qy 379 ITTHPESNQGLITTRKGLDFKXNQHTLYVEVTNEAPVLKI-----PTSTATIVVHVE 432
Db 501 LDWEPDSGHVRLRLCKNLSEAPASHVEVVVVQS-----VAKLVGPGPGGATATVTVLVE 556

Qy 433 DVNEAPVFPKSVKVE--VQSGIPTGPV--CVVTAEDPPKENQKISVRIILRDPAGLAM 488
Db 557 RV-----MPPPKLDQSEYASAPAGSFLTLTIQPSDPISTRFLPSLVNDSEGLICI 610

Qy 489 DPDSQVTAAGTLDREDDQFVRNNIYVWVLAMDNNGSPPTTGTGTGLTLILLIDVNDHGFPV 548

Db 611 EKFSGEVHTAQSLQGAQP-----GDTYTVLVVEAQDTA-----LTL-----APVP 649

Qy 549 E-----PRQITICNQSPVRHLNIT-----DKDLSPHTSPFOAQLTDDSDIY--WTAEVNE 597

Db 650 SQYLCTPRQ-----DHGLIVSGSPKDPDLSASGHGYSFTLGNPTVQDRWLQTLN 700

Qy 598 EGDVTVLSLKFKLQDQTYDVHLSLSDHGKNEQLTVIRATVCDCH--GHVETCPGPMKG-- 653

Db 701 GSHAYLTALHWHVEPREHIIPVVVSHNAQMQL-LVRVIVCRNVEGQCQMKRKGMP 759

Qy 654 -----GFILPVLGAVLALLFLLVLLVLRKK 680

Db 760 TKLSAVGILVGTILVAIGIFILIFLIFTHWTMSRKK 792

RESULT 12
US-10-131-826A-394
Sequence 394, Application US/10131826A
Publication No. US20050245730A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
APPLICANT: Watanabe, Colin K
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: F3330RIC128
CURRENT APPLICATION NUMBER: US/10/131,826A
CURRENT FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 394
LENGTH: 1184
TYPE: PRT
ORGANISM: Homo Sapien
US-10-131-826A-394

Query Match 7.6%; Score 333.5; DB 6; Length 1184;
Best Local Similarity 21.1%; Pred. No. 1.5e-17;

Matches 191; Conservative 139; Mismatches 378; Indels 199; Gaps 36;	
QY	6 GPLASLLLVQVW-----LQCAASEPCRAVREAEVTLAAGGAEQEPGQALGVPMG 57
Db	13 GPGGYLLGDCQEVTTLTVKYQVSEBVPSTGTVIGKLSQEL---GREERRQA-GAAFOV 68
QY	58 CPQEPALFSTDNDDFTVRNGETVQERRSLKERNPLKIFPSKILRRHKRDWVAISVP 117
Db	69 LQLPQALPIQVDSSEGLSTGRDLRQCLRWDPCLV--SFDVLTATGDIALHVEIQVL 126
QY	118 E-NGKGP-PPQRLNQLKSNKDRDTKIFYSI-----TGPGA---DSPPEGVFAV--- 160
Db	127 DINDHQPRFKPQGELEISSASLRTPIPLDRALDPDTGENTLHTLSPSEHFALDVIV 186
QY	161 ---EKETGMLLNKPLDREIEAKYELFGHVAENGASVEDPMNISIIITDQNDHKPKFTQ 217
Db	187 GPDTKHAELIUVKELDREIHSFFDLVLTAYDNGNPKSGTSLVKVNLDSNDNSPAFAE 246
QY	218 DTRGSLVLEGVLPGTSVMQVATDEDDAIYTYNGVAVSHSQEPKDPDPLDMFTIHRSTG 277
Db	247 SSLALEIQEDAAPGTLILKLTATDPDQG---PNGEVEFFLTKMPPEVLD-TFSIDAKTG 302
QY	278 TISVISGLDREKVPVETLTQATDMGDGSGTTTAVAVVEILDANDNAPMF-----DPQ 331
Db	303 QV-ILRPLDYENKPAVEVDVQARDLGNPIPAHCKVKILKVLVDVNDNISIHVTWASQPS 361
QY	332 KYBAHPENAVGHEVORLTVTDLDPNS---PAWRATYILMGDDGDGHFTTITTHPSNQ 388
Db	362 LVSEALPKDSF---IALWADDLDSGHGLVHCLSQEL-----GHFRL-KRTNGNTY 410
QY	389 ILTRKGLDPEAKNQHTLYVEVNEAPFVLKLTSTATIVVHVEDNE-APVPVPSKV 447
Db	411 MLLTNATLDREQWPKYTLTLAQDQG---LQPLSAKKQLSIQISDINDNAPVEKRYEV 467
QY	448 EVOEGITGPVCVYTAEDPKENOKISYRILRDPAGWL-AMDPDSGQVTAAGTLDRED 505
Db	468 STRENNLPSLHTITIKADHDLGINKVSTRIQDSPVAHLVAIDSNTEVTAORSNYEE 527
QY	506 EQFVRNNIYEMVVLAMDNQSPPTGTGTLTLTLDVNDHGP-VPEP-----RQITICNQ 558
Db	528 MAG-----PEFQVIAEDSGQPLASSVSVWVSLLDANDNAPEVQVPLSDGKASLVN 582
QY	559 SPVRHVL-----NITDKD----- 571
Db	583 ASTCHLLVPIETPNGLGPGAGTTPPLATHSSRPFLTTIIVARDADSGANGEPLYSIRNG 642
QY	572 -----LSPHTSPFQQLTDDSDIYWTAEVNEGDTVLSLKKFLKQDTVDVHLSLSDHG 625
Db	643 EAHFLINPHTGOLFVNVTNASSL-----IGSE-----WELBIVVEDQG 681
QY	626 N-----KEQLTVIRATVCDCHGHVETCPGPWKGGFILPVLGAVLALLF-LLLVLLLVK 679
Db	682 SPPLQTRALLRVMTFVSDHLRDSARKPGALSMLTVCILAVLLGIFGLILALFMSICR 741
QY	680 KRKIKPLLPEDDTRNDNPFYGEGBEDQDYDTQLHRGLEARPEVVVLNRDVAPTII 739
Db	742 TEK-----KDNRAYNCREA-----ESTYRQOPKRPQKHQK---ADIH 776
QY	740 PTWYRPRPNDPEIG---NFIIENLKAANTDP-----TAPPYDTLLVFDYEGSGS 787
Db	777 LVPVLRQAGEPCVEVGQSHKVDKEAMMEAGWDPCLOAPHLPTLYRTL-----RNQGN 831
QY	788 DAASLSS 794
Db	832 QGAPAES 838

RESULT 13
US-10-131-826A-130
; Sequence 130, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen	
; APPLICANT: DeForge, Laura	
; APPLICANT: Desnoyers, Luc	
; APPLICANT: Filvaroff, Ellen	
; APPLICANT: Gao, Wei-Qiang	
; APPLICANT: Gerritsen, Mary E.	
; APPLICANT: Goddard, Audrey	
; APPLICANT: Godowski, Paul J.	
; APPLICANT: Gurney, Austin L.	
; APPLICANT: Sherwood, Steven	
; APPLICANT: Smith, Victoria	
; APPLICANT: Stewart, Timothy A.	
; APPLICANT: Tumas, Daniel	
; APPLICANT: Watanabe, Colin K	
; APPLICANT: Wood, William	
; APPLICANT: Zhang, Zemin	
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC	
; FILE OF INVENTION: ACIDS ENCODING THE SAME	
; FILE REFERENCE: P3330R1C128	
; CURRENT APPLICATION NUMBER: US/10/131,826A	
; CURRENT FILING DATE: 2002-04-24	
; PRIOR APPLICATION NUMBER: 60/049911	
; PRIOR FILING DATE: 1997-06-18	
; PRIOR APPLICATION NUMBER: 60/056974	
; PRIOR FILING DATE: 1997-08-26	
; PRIOR APPLICATION NUMBER: 60/059113	
; PRIOR FILING DATE: 1997-09-17	
; PRIOR APPLICATION NUMBER: 60/059115	
; PRIOR FILING DATE: 1997-09-17	
; PRIOR APPLICATION NUMBER: 60/059117	
; PRIOR FILING DATE: 1997-09-17	
; PRIOR APPLICATION NUMBER: 60/059122	
; PRIOR FILING DATE: 1997-09-17	
; PRIOR APPLICATION NUMBER: 60/059184	
; PRIOR FILING DATE: 1997-09-17	
; PRIOR APPLICATION NUMBER: 60/059263	
; PRIOR FILING DATE: 1997-09-18	
; PRIOR APPLICATION NUMBER: 60/059352	
; PRIOR FILING DATE: 1997-09-19	
; PRIOR APPLICATION NUMBER: 60/059588	
; PRIOR FILING DATE: 1997-09-19	
; Remaining Prior Application data removed - See File wrapper or PALM.	
; NUMBER OF SEQ ID NOS: 550	
; SEQ ID NO 130	
; TYPE: PRT	
; LENGTH: 530	
; ORGANISM: Homo Sapien	
US-10-131-826A-130	

Query Match 6.4%; Score 280; DB 6; Length 530;	
Best Local Similarity 27.2%; Pred. No. 5.3e-14;	
Matches 126; Conservative 65; Mismatches 200; Indels 72; Gaps 23;	
QY	208 QNDHRPKFTQDTRGSLV---EGVLPGTSVMQVATDEDDAIYTYNGVAVSHSQEPKDP 265
Db	24 QVNRLPFFTHFFDFTVLLISEDTPVGSSVTQLLAQDMD-----NDPLVFGVSGEEA--- 74
QY	266 HDLMFTIHRSTGTISVISSGLDREKVPVETLTQATDMGDGSGTTTAVAVVEILDANDNA 325
Db	75 -SRFAVEPDTGVV-WLRQPLDRETKSEFTVFSVDHQG---VITRKNIQVGDVNDNA 129
QY	326 PMFDPQKYBAHPENA-VGHEVQRLTVTDLDPNSPAWRATYILMGCD-----DGDHF 377
Db	130 PTFHNPQSVRIPEPTPVGTPIFVNATDPD-----LGAGGSVLVSFPSPSQFF 178
QY	378 TITTHPESNOGLITRKGDLDFEAKQOHTLYVEVNEAPFVLKLTPTST-----ATIVVHVED 433
Db	179 AI-----DSARGIVTVIRELDYETTOAYQLTVNATDQDK---TRPLSTLANLAIIITDQD 231
QY	434 VNEAPVFPVPSKVVEQEGIPTEGPEVCVVTAEADPKENOK-ISYRILRDPAGWL-AMDPD 491
Db	232 MD--PIFINLPYSTNIYEHSPPTVRIITAITDQDKRPRGIGYTIUSGNTNSIFALDYI 289

Db 1252 RTVTAFVEKDEG-----MLEEAEUPRASDRNIMCEDDCHDITYYSIVGNGSGHFTV--D 1305

Qy 383 PESNOGILTRKGLDFEAKNQHTLYVEVTNEAPFVLKLPSTSTATIVVHVVEDVNEAPVFPV 442

Db 1306 PRTN--VLSLVKPLDRSEQETHLIIGASDTNPAAVLQASTLTVTNVVREANPRPVF-- 1361

Qy 443 PSKVVEQEGIPTEGPV--CVYTAEDDDKENQKISYRILRDPAGWLAMDPSQGVTA VGT 500

Db 1362 --QRALYTAGISAGDFTIERNLLTLVATHSEDLPTTYTIOE-----SMEADP-----T 1407

Qy 501 LDR-EDQEQVRN-----NIVEVMVLAMDNGSPPTTGCTGLLTLIDVN 542

Db 1408 LEAVQSAFILNPETGVLSNFQPTASMGHGFVFAVADSRT-ETARTEVKVILISDRN 1466

Qy 543 D-----HGPVPE--PROITI-----CN-----QSPVRHVLNITDKDLSPHTS 577

Db 1467 RVFFTFNNPLPEVTPQEDFIAETFTAFFGTCNIDQSWASDDPVTGATKDDQTEVRAHF- 1525

Qy 578 PFOAQLTDDSDIYWTAEVNEE---GDTVLSLKKFLKQDTYDVHLSLSDHGKNEQLTVIR 634

Db 1526 -----IRDDLFPV--PAEEIEQLRGNPTLVNSIQRALEEQ-----NLQLAD----- 1563

Qy 635 ATVCDCHGHVETCPGPKGC-----FILPVLGAVLALFLLLVLLLRKK--RKIK 684

Db 1564 -----LFTGETPILGGDAQARALYALAAVALIVVLLIVFFVTRTTLNRRL- 1613

Qy 685 EPLLPEDDTRDN-----VFYEGEGGGE-----EDQDYDITQLH 719

Db 1614 QALSMTKYSSQDSGLNRVGLAAPGTNKHAVEGSPINWETLKAPDFDALSEQSYDSDLI- 1672

Qy 720 RGLARPVVLRNDVAPTIPTMYRPRAN-----PDEIGNFIENLKAANTDP 769

Db 1673 -GIEDLPQ--FRNDYFP-----PEGSSMRGVVNEHVPESIANH--NNNFGFNSTP 1718

Qy 770 TAPPY 774

Db 1719 FSPEF 1723

Search completed: December 12, 2005, 09:20:10
Job time : 15 secs

THIS PAGE BLANK (USPTO)